PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12N 15/11, 15/00, 15/63, C07H 21/02, 21/04

(11) International Publication Number:

WO 00/04140

(43) International Publication Date:

27-January 2000 (27.01.00)

(21) International Application Number:

PCT/US99/15849

(22) International Filing Date:

14 July 1999 (14.07.99)

(30) Priority Data:

60/092,921 15 July 1998 (15.07.98) US 60/092,922 15 July 1998 (15.07.98) US 60/092,956 15 July 1998 (15.07.98) US

(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]: 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). KOMATSOULIS, George [US/US]; 9518 Garwood Street, Silver Spring, MD 20901 (US). DUAN, Roxanne, D. [US/US]; 5515 Northfield Road, Bethesda, MD 20817 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US). SHI, Yang-gu [CN/US]; Apartment 102, 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 3142 Quesada Street, N.W., Washington, DC 20015 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne

Terrace #316, Gaithersburg, MD 20878 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; Apartment 115, 410 Van Dyke Street, St. Paul, MN 55119-4321 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). MUCENSKI, Michael [US/US]; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US).

(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: 71 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia .
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
ВJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		•
EE	Estonia	LR	Liberia	SG	Singapore		

71 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

5

10

15

20

25

30

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human

2

growth hormone, tissue plasminogen activator, and erythropoeitin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

15

10

5

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

20

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

25

30

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing

3

to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

In specific embodiments, the polynucleotides of the invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of the coding sequence, but do not comprise all or a portion of any intron. In another embodiment, the nucleic acid comprising the coding sequence does not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene in the genome).

5

10

15

20

25

30

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an

4

overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μ g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

5

10

15

20

25

30

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of

5

single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

5

10

15

20

25

30

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation,

5

10

15

20

gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 1

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

30 PFCSGFFPSLWIYLPFIFNVSDLWMGSLSGCALPFCLXVFFLTVSPSAVGLLXF AGGPLQTLFAWVSPVEAAEQQRLLPVLSSGSFVSEGTCQMPARALLYEVSVG

7

PYWEIPPSQDTRRSGTYLRRQSDP (SEQ ID NO: 195). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in pancreas islet cell tumors.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the pancreas, including cancer and diabetes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pancreas, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endocrine, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tumors of pancreatic islet cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment and intervention of such tumors, in addition to other endocrine or gastrointestinal tumors where expression has been indicated. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

8

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1099 of SEQ ID NO:11, b is an integer of 15 to 1113, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

5

10

15

20

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

HEGSCRAPGFSAHKGRGCPSPRMTLPSRALASLGVGVWGMLRLNQVTVSCG GSRWSSRVALGAFSWVCGVALVLQPSGGGLGLTSPSEGCWEGELALAVLRA PGGSPS (SEQ ID NO: 196) . Polynucleotides encoding these polypeptides are also provided.

This gene is expressed equally in in.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic disorders, particularly leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hemolymphoid, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

9

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 104 as residues: Gly-29 to Ser-35, Ser-63 to Cys-68. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution in hemangiopericytoma, breast lymph node, and bone marrow indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 969 of SEQ ID NO:12, b is an

10

integer of 15 to 983, where both a and \ddot{b} correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where \ddot{b} is greater than or equal to \ddot{a} + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

5

10

15

20

25

30

The translation product of this gene shares sequence homology with the Drosophila melanogaster slit protein, a secreted protein that contains both an EGF domain and Leucine Rich Repeat domains. It is thought to be important in the development of midline glia and commissural axon pathways (See e.g., Rothberg et al. Genes Dev. 4:2169-87 (1990); which is hereby incorporated by reference herein).

This gene is expressed primarily in human hippocampus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neurological, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution within human hippocampus combined with the homology to the Drosophila slit protein, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease,

11

Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

5

10

15

20

25

30

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 959 of SEQ ID NO:13, b is an integer of 15 to 973, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by

12

the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

IPLTLPGIFLLIRLFWRLGQSICGPGKLVLWPQFCCGCAVISGHCVPRGMPSSW LPGCFVLLCLVAVGCQLREWGVGGVSAVGLLALPHLQVLGMRGRGLISGG (SEQ ID NO: 197). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 16. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 16.

This gene is expressed in KMH2 cells, osteoblasts, fetal spleen, Jurkat membrane bound polysomes, breast, and cerebellum.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer, immune, and skeletal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, skeletal, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in KMH2 cells, osteoblasts, and fetal spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in fetal spleen and T-cells indicates a role in the regulation of the proliferation: survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved

in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1444 of SEQ ID NO:14, b is an

14

integer of 15 to 1458, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

5

10

15

20

25

30

The translation product of this gene shares sequence homology with phospholipase A2 which cleaves fatty acids from carbon 2 of glycerol (ref. Prosite pattern documentation for PS2_HIS). Many snake venoms contain phospolipase A2, which prevents transmission of nerve impulses to muscles by blocking the release of acetylcholine from the neuron. Therefore, included in this invention as preferred domains are Phospholipase A2 histidine active site domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). Phospholipase A2 is an enzyme which releases fatty acids from the second carbon group of glycerol. Structurally, PA2's are small and rigid proteins of 120 amino-acid residues that have four to seven disulfide bonds. PA2 binds a calcium ion which is required for activity. The side chains of two conserved residues, a histidine and an aspartic acid, participate in a 'catalytic network'. Two different signature patterns for PA2's were developed. The first is centered on the active site histidine and contains three cysteines involved in disulfide bonds. The consensus pattern is as follows: C-C-x(2)-H-x(2)-C [H is the active site residue].

Preferred polypeptides of the invention comprise a Phospholipase A2 histidine active site domain selected from the following amino acid sequences: CCNQHDRC (SEQ ID NO: 199), SLTKCCNQHDRCYET (SEQ ID NO: 200), and/or LTKCCNQHDRCYETCG (SEQ ID NO: 201). Polynucleotides encoding these polypeptides are also provided. Further preferred are polypeptides comprising the Phospholipase A2 histidine active site domain of the sequence listed in Table 1 for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C- terminal to the Phospholipase A2 histidine active site domain. Alternatively, the additional contiguous amino acid residues is both N-terminal and C-terminal to the Phospholipase A2 histidine active site domain, wherein the total N-and C-terminal contiguous amino acid residues equal the specified number. The

above preferred polypeptide domain is characteristic of a signature specific to Phospholipase A2 proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with Phospholipase A2 proteins. Such activities are known in the art, some of which are described elsewhere herein, or see, for example, McIntosh, et al. J. Biol. Chem. 270 (8), 3518-3526 (1995), incorporated herein by reference.

5

10

15

20

25

30

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

GPAGKEAWIWSWLLPSPGPAPLPSASWGLCGDAPR

AAARGPVEPGAARMALLSRPALTLLLLLMAAVVRCQEQAQTTDWRATLKTI
RNGVHKIDTYLNAALDLLGGEDGLCQYKCSDGSKPFPRYGYKPSPPNGCGSP
LFGXHLNIGIPSLTKCCNQHDRCYETCGKSKNDCDEEFQYCLSKICRDVQKTL
GLTQHVQACETTVELLFDSVIHLGCKPYLDSQRAACRCHYEEKTDL (SEQ ID
NO: 198). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed in a variety of cell types with no single type predominating.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological disorders, or metabolism disorders, specifically phospholipase A2 deficiencies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neuromuscular system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., pancreas, cancerous and wounded tissues) or bodily fluids (e.g., bile, lymph,

16

serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic

5

10

20

25

30

above listed tissues.

epitopes shown in SEQ ID NO: 107 as residues: Gln-23 to Asp-30, Lys-66 to Cys-87. Polynucleotides encoding said polypeptides are also provided. The ubiquitous tissue distribution and homology to phospholipase A2 indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neuromuscular disorders. Alternatively, considering the activity of phospholipase A2 as a block for neuro-transmission may suggest that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the homology to Phospholipase A2 proteins may indicate a potential use for the protein product of this gene in diagnosis, treatment and/or prevention of metabolism disorders, specifically deficiencies in Phospholipase A2. Furthermore, the protein may also be used to

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of

determine biological activity, to raise antibodies, as tissue markers, to isolate cognate

ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the

protein may show utility as a tumor marker and/or immunotherapy targets for the

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1991 of SEQ ID NO:15, b is an integer of 15 to 2005, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

5

10

15

20

25

30

In a specific enbodiment polypeptides of the invention comprise the following amino acid sequence:

GTSSARPRGALPGGSAPSAPHGQLPGRAQPAPVSGPPPTSGLCHFDPAAPWPL WPGPWQLPPHPQDWPAHPDIPQDWVSFLRSFGQLTLCPRNGTVTGKWRGSH VVGLLTTLNFGDGPDRNKTRTFQATVLGSQMGLKGSSAGQLVLITARVTTER TAGTCLYFSAVPGILPSSQPPISCSEEGAGNATLSPRMGEECVSVWSHEGLVLT KLLTSEELALCGSRLLVLGSFLLLFCGLLCCVTAMCFHPRRESHWSRTRL (SEQ ID NO: 202) . Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

ARAPPGPEGLSPEAQPPLLPMGNCQAGHNLHLCLAHHPPLVCATLILLLLGLS GLGLGSFLLTHRTGLRT LTSPRTGSLF (SEQ ID NO: 203) . Polynucleotides encoding these polypeptides are also provided.

This gene is expressed in a wide variety of tissue types including testes, cerebellum, dendritic cells, breast cancer, umbilical vein endothelial cells, epididymus, corpus colosum, chronic synovitis, liver hepatome, normal breast, osteoblasts, melanocytes, B cell lymphomas, and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, cancer, particularly of endothelial tissues. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endothelial, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 108 as residues: Thr-52 to Gly-57. Polynucleotides encoding said polypeptides are also provided.

Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that the protein product of this gene may play a role in the regulation of cellular division and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also relies on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have

19

applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 929 of SEQ ID NO:16, b is an integer of 15 to 943, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

RFLSVXPQXEVPFLLHPCVCFXGGHPSLLPDPCRAVGGGWEAPRCCLHEALC QSLGCKAEEIVSVSESSSAQRCWYLLRGRKAGGRGPASPVLFALMRLESLCH LCLACLFFRLPATRTVYCMNEAEIVDVALGILIESRKQXKACEQPALAGADNP EHSPPCSVSPHTSSGSSSEEEDSGKQALXPGLSPSQRPGGSSSACSRSPEEEE EEDVLKYVREIFFS (SEQ ID NO: 204) . Polynucleotides encoding these polypeptides are also provided. Polynucleotides of the invention do not consist of the nucleic acid sequences shown as GeneSeq Accession Nos: V59595 and V59744, which are hereby incorporated herein by reference.

5

10

15

20

25

30

This gene is expressed primarily in a variety of immune cell types, including stromal cells, dendritic cells, leukocytes, activated T-cells, macrophages, monoctyes, neutrophils and to a lesser extent in a variety of other adult and fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in fetal tissue and various hematopoietic cancers indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all

21

hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1489 of SEQ ID NO:17, b is an

22

integer of 15 to 1503, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

5

10

15

20

25

30

When tested against Jurkat T-cell lines, supernatants removed from cells containing this gene activated the NF-kB (Nuclear Factor kB) pathway. Thus, it is likely that this gene activates T-cells through the NF-kB signal transduction pathway. NF-kB is a transcription factor activated by a wide variety of agents, leading to cell activation, differentiation, or apoptosis. Reporter constructs utilizing the NF-kB promoter element are used to screen supernatants for such activity. Preferred polypeptides of the invention comprise the following amino acid sequence: VPGWPRACSPCQADSPRAHPPKLRGILRWAPVPLXCAALCPPLDSG MSMAACPEAPEPSFLREVPSSPASTQWHRPCNFRQVEANPRKEPKNLVWRD VSLGQXSRTPRGSGLELVRVCGGGMQRDKTVVEERVGEERERERESLGG AGKHGEMRCVYVRESVGAPGRAGGGGNGVNSVGCVRTVHSGSXPPPSAGV S (SEQ ID NO: 205). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in parts of the brain such as cerebellum and frontal lobe.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

23

The tissue distribution in cerebellum and frontal lobe indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, prevention and/or treatment of neurodegenerative disease states and behavioural disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

10

15

20

25

30

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1498 of SEQ ID NO:18, b is an

24

integer of 15 to 1512, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

5

10

15

20

25

30

In a specific embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

TRPGKELNLVFGLQLSMARIGSTVNMNLMGWLYSKIEALLGSAGHTTLGITL MIGGITCILSLICALALAYLDQRAERILHKEQGKTGEVIKLTDVKDFSLPLWLIF IICVCYYVAVFPFIGLGKVFFTEKFGFSSQAASAINSVVYVISAPMSPVFGLLV DKTGKNIIWVLCA (SEQ ID NO: 206) . Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in fetal tissue, and to a lesser extent in a variety of adult human tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fetal abnormalities, particularly developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 111 as residues: Lys-30 to Thr-35. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in fetal tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Furthermore, the protein

26

may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1641 of SEQ ID NO:19, b is an integer of 15 to 1655, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

5

10

15

20

25

30

The translation product of this gene shares sequence homology with human histiocyte-secreted factor (HSF) which is a novel cytokine that shows in vivo antitumour activity without the cytotoxicity associated with tumour necrosis factor. Furthermore, The translation product of this gene also shares sequence homology with the human endogenous virus S71 gag polyprotein, the sequence of which is believed to represent a transformation locus for several cancers (See Genebank Accession No. pir|A46312|A46312; all references available through this accession are hereby incorporated by reference herein). Similarly, The translation product of this gene also shares homology with B219, a sequence that is expressed in at least four isoforms in very primitive hematopoietic cell populations which may represent a novel hemopoietin receptor (See, e.g., Cioffi, et al. Nat. Med. 2:585-589 (1996), which is hereby incorporated by reference herein). In a preferred embodiment polypeptides of the invention comprise the following amino acid sequence:

27

CKDLCSRVYLLTLSPLLSYDPATSHSPRNTQ (SEQ ID NO: 207). Also preferred are the polynucleotides encoding these polypeptides.

This gene is expressed primarily in tonsil, and colon, and to a lesser extent in a wide variety of human tissues.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, and gastrointestinal disorders, particularly tumors of the colon and tonsil. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic, digestive and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, gastrointestinal, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 112 as residues: Met-1 to Cys-6. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in tonsil and colon, combined with the homology to human histiocyte growth factor, the human endogenous viral protein, and B219 strongly indicate that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment and/or prevention, of a variety of hematopoietic and immune system disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation

28

of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2511 of SEQ ID NO:20, b is an

29

integer of 15 to 2525, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

5

10

15

20

25

30

The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

IICECWEEECQSCRLKITQPREICRMDFLVLFLFYLASVLMGLVLICVCSKTHS LKGLARGGAQIFSCIIPECLQRAXHGLLHYLFHTRNHTFIVLHLVLQGMVYTE YTWEVFGYCQELELSLHYLLLPYLLLGVNLFFFTLTCGTNPGIITKANELLFLH VYEFDEVMFPKNVRCSTCDLRKPARSKHCSVCNWCVHRFDHHCVWVNNCI GAWNIRYFLIYVLTLTASAATVAIVSTTFLVHLVVMSDLYQETYIDDLGHLHV MDTVFLIQYLFLTFPRIVFMLGFVVVLSFLLGGYLLFVLYLAATNQTTNEWYR GDWAWCQRCPLVAWPPSAEPQVHRNIHSHGLRSNLQEIFLPAFPCHERKKQE (SEQ ID NO: 208) . Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in colon and brain and to some extent in all tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and digestive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and digestive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neurological, gastrointestinal, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Alternatively, expression of this gene in colon may indicate a role in the detection, prevention and/or treatment of colon diorders such as colon cancer, Crohn's Disease, ulcers, and digestive tract disorders in general. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of

31

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1382 of SEQ ID NO:21. b is an integer of 15 to 1396, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

5

10

15

20

25

30

When tested against Reh cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activation site) pathway. Thus, it is likely that this gene activates B-cells through the Jaks-STAT signal transduction pathway. GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. This gene maps to chromosome 7, and therefore, is used as a marker in linkage analysis for chromosome 7.

This gene is expressed primarily in brain, and in the developing embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, behavioral, immune, and developmental disorders.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and developmental systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developing, immune, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue

32

or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 114 as residues: Lys-60 to Asn-67. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the tissue distribution in developing embryo indicates that the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the biological activity within B-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Activation of genes witin B-cells indicates a role for this protein in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

33

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1055 of SEQ ID NO:22, b is an integer of 15 to 1069, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

20

30

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

The gene encoding the disclosed cDNA is believed to reside on chromosome 6. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 6.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

following amino acid sequence:

LLSFKIRGLRTEDAGWAQSSSGGLCVRGDAFWMPSSSSGLGSPSRPPSSFLCL LLLLLPPAALALLLFFLDFFPPRAAVSPFLPDHCSARQPRVWRRETLNRSASGL

34

GCWARSTEQGAVGVATGTVLDI SLPASCLSLWPPGPSGGI (SEQ ID NO: 209) . Polynucleotides encoding these polypeptides are also provided.

In a specific embodiment polypeptides of the invention comprise the following amino acid sequence:

5 QLGLCLTSASLPPASRCGHQAPLGASDLSAHHSAPGFSDSYFTMSCQSSLSRA
EILQCPLVPSVSPPTHLPQGRANKSSRASLPLLPQTHWCLFPSARGWRRGIQSG
LPPGGSCTSPRSPPQTLHQHITLVNHNTSYWQSPST (SEQ ID NO: 210),
HQPPCLLPLAVATRPLWGHLTCLPIILHLVSVTLTSPCLANQAFQGQRSYNAL
WCPLFLLLPTSPKGEQTNHPEPACPCFPKLTGVFSLQHVVGAEEFSQVFLLVD
PVPVLDHLLKLFTSTSHLLIIIPHIGKAPAPDSLL EELSLSLATHCKVAVARFT
(SEQ ID NO: 211). Also preferred are the polynucleotides encoding these
polypeptides. Polynucleotides of the invention do not consist of the nucleic acid
sequence shown as GeneSeq Accession No. X04377, which is hereby incorporated
herein by reference.

This gene is expressed primarily in brain.

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, behavioral and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 115 as residues: Pro-2 to Gly-7, Ser-10 to Ser-16,

35

Pro-52 to Val-62, Arg-64 to Ser-73. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1644 of SEQ ID NO:23, b is an

36

integer of 15 to 1658, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

5

10

15

20

25

30

The translation product of this gene was shown to have homology to the lysosomal mannosidase alpha-B protein (See Genebank Accession No. P34098) which is thought to be important in protein metabolism. One embodiment of this gene comprises polypeptides of the following amino acid sequence:

MAAEGSRFSSQSPGLVDRQGPKCDPSRLVSPWGRHGLRILQIGHHHGRDGQH EATHHLLRVLRAPRVGKADEGAVDSDPSTPLQLKHEAAHAEDHAQQVHVVR RRVVQGRVTFARRGLVPQHFVRPPWVRHIVSGHSESKARSRLFRCRNRSFRR AS (SEQ ID NO: 212), and/or

RLVSPWGRHGLRILQIGHHHGRDGQHEATHHLL RVLRA (SEQ ID NO: 213). An additional embodiment is the polynucleotides encoding these polypeptides. This gene maps to chromosome 19, and therefore, is used as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in brain, placenta, fetal liver, and to a lesser extent in most tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, reproductive, and hepatic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hepatic, or cancerous and wounded tissues) or bodily fluids (e.g., bile, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 116 as residues: Asn-34 to Lys-42, Leu-60 to Trp-70. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution predominantly in brain indicates a role in the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntinton's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Alternatively, the tissue distribution in liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-healing models and/or tissue trauma.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1063 of SEQ ID NO:24, b is an integer of 15 to 1077, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

25

30

5

10

15

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in spinal cord, Merkel cells, and adipose tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

38

not limited to, disorders of the nervous and immune systems, particularly those disorders relating to the CNS involving lipid metabolism disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and immune systems and adipose tissue, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, immune, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

The tissue distribution in spinal cord, Merkel cells and adipose tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of diseases the nervous systems, such as spinal cord injury, neurodegenerative diseases, muscular dystrophy or obesity. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1191 of SEQ ID NO:25, b is an integer of 15 to 1205, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

The translation product of this gene shares sequence homology with the human uncoupling protein-2 which is thought to be important in energy metabolism, obesity, and the predisposition of hyperinsulinemia (See Genebank Accesion No. gi|1857278). Recently, another group published on this gene, designating it brain mitochondrial carrier protein-1 (BCMP1) (J Biol Chem 1998 Dec 18;273(51):34611-5). One embodiment of this gene comprises polypeptides of the following amino acid sequence: PTDVLKIRMQAQ (SEQ ID NO: 214), and/or TYEQLKR (SEQ ID NO: 215). An additional embodiment is the polynucleotides encoding these polypeptides. This gene maps to the X chromosome, and therefore, is used as a marker in linkage analysis for the X chromosome.

5

10

15

20

25

30

This gene is expressed primarily in manic depression brain tissue, epileptic frontal cortex, human erythroleukemia cell line, T-helper cells, and to a lesser extent in endothelial and amygdala cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system or hematopoietic/immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system or hematopoietic/immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hemolymphoid, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 118 as residues: Ser-34 to Thr-39, Gln-198 to Leu-205. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neural tissues combined with the homology to the human uncoupling protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and/or treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Alternatively, given the homology to uncoupling proteins, the gene and/or its translation product may also be used in the diagnosis, treatment, and/or prevention of thermogenesis disorders such as obesity, cachexia, and hyperinsulinemia. Uncoupling proteins dissipate the proton gradient created from the oxidation of fuels by the electron transport chain, thus releasing stored energy as heat. Dysfunction of thermogenesis can induce disorders such as obesity and cachexia. It is thought that obesity may result from decreased thermogenesis in humans. Alternatively, cachexia is a metabolic state in which energy expenditure exceeds food intake, for example in anorexia nervosa. Uncoupling proteins is useful for the treatment and/or prevention of diseases and/or disorders involved with aberrant metabolic and thermogenic pathways. The following method provides for the determination of respiration uncoupling activity of the polypeptides of the present invention, including fragments and variants of the full length proteins.

Briefly, yeast are transfected with an expression vector expressing polypeptide of the present invention as previously described by Bouillaud et al., EMBO J., 13:1990 (1994) (incorporated by reference herein in its entirety). Rates of growth in liquid medium of transformed yeast are measured in the presence of galactose, which induces expression, as described in International Publication No. WO 98/31396

5

10

15

(incorporated by reference herein in its entirety). Instanteous generation times are compared between the polypeptide of the present invention and appropriate controls. An in vivo decrease of membrane potential associated with uncoupling of respiration is analyzed by flow cytometry of yeast labeled with the potential sensitive probe DiOC6 (3) (3,3'-dihexyloxacarbocyanine iodine, Molecular Probes, Eugene, OR). The ability of a polypeptide of the present invention to influence mitochondrial activity and uncouple respiration is thus determined.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1660 of SEQ ID NO:26, b is an integer of 15 to 1674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

The translation product of this gene shares sequence homology with 55 kD deglycosylated zona pellucida protein which is known to be important in egg fertilization (See Genebank Accession No.R39356). Preferred polypeptides of the invention comprise the following amino acid sequence:

25 RPRPSASSLARSASLLPAAHGXGVGGAGGGSSXLRSRYQQLQNEEESGEPEQ
AAGDAPPPYSSISAESAHXFDYKDESGFPKPPSYNVATTLPSYDEAERTKAEA
TIPLVPGRDEDFVGRDDFDDADQLRIGNDGIF (SEQ ID NO: 216),
RYQQLQNEEESGEPEQAAGD (SEQ ID NO: 217), and/or
PGRDEDFVGRDDFDDADQLRIG (SEQ ID NO: 218). Polynucleotides encoding
30 these polypeptides are also provided.

42

Preferred polypeptide fragments of the invention comprise the following amino acid sequence: MLTFFMAFLFNWIGFFLSFCLTTSAAGRYG AISGFGLSLIKWILIVRFSTYFPGYFDGQY WLWWVFLVLGFLLFLRGFINYAKVRKMPET FSNLPRTRVLFIY (SEQ ID NO: 219). Polynucleotides encoding these polypeptides are also provided.

Preferred polypeptide varients of the invention comprise the following amino acid sequence:

5

10

15

20

25

30

RNNHLL (SEQ ID NO: 222), and/or

MKKSLENLNRLQVMLLHLTAAFLQRAQHXFDYKDESGFPKPPSYNVATTLPS
YDEAERTKAEATIPLVPGRDEDFVGRDDFDDADQLRIGNDGIFMLTFFMAFLF
NWIGFFLSFCLTTSAAGRYGAISGFGLSLIKWILIVRFSTYFPGYFDGQYWLW
WVFLVLGFLLFLRGFINYAKVR KMPETFSNLPRTRVLFIY (SEQ ID NO: 220),
MLLHLTAAFLQRAQFSTYFPGYFDGQYWLWWVFLVLGFLLFLRGFINYAKV
RKMPETFSN LPRTRVLFIY (SEQ ID NO: 221), MLTFFMAFLFNWIGFFLSFCLT
TSAAGRYGAISGFGLSLIKWILIVRFSTYFPAFMNSLSRSKRTPAGSESRCRTQ

MKKSLENLNRLQVMLLHLTAAFLQRAHXIL TTRMSLGFQSPHLTM (SEQ ID NO: 223) . Polynucleotides encoding these polypeptides are also provided.

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in adult kidney, colon adenocarcinoma, and fetal brain, and to a lesser extent, ubiquitously expression in many tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, disorders of kidney, colon cancers, and central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal and neural systems, and cancers, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., renal, neural, urogenital, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 119 as residues: Cys-15 to Gly-36. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution adult kidney, colon adenocarcinoma, and fetal brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of kidney diseases, colon cancers, and disorders of the central nervous system. Additionally, the homology to the zona pellucida protein indicates that the gene product is used for male contraceptive development, and infertility diagnosis etc. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1951 of SEQ ID NO:27, b is an

44

integer of 15 to 1965, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

5

10

15

20

25

30

The translation product of this gene shares sequence homology with the chicken transferrin receptor in addition to a human prostate-specific protein homolog (See Genebank Accession Nos.pir|JH0570|JH0570 and gi|2565338 (AF026380), respectively). This gene also shares significant homology with both the murine and the rat hematopoietic lineage switch 2 proteins (See Genbank Accession Nos. g3169729 and g3851632, respectively), which are induced during an erythroid to myeloid lineage switch.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MTVMDPKQMNVAAAVWAVVSYVVADMEEML PRS (SEQ ID NO: 224). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in fetal tissues, such as liver/spleen and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pre-natal disorders, anomalies, deficiencies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing fetus, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 120 as residues: Arg-31 to Lys-37, Lys-58 to Glu-65,

Asp-157 to Gly-168, Ile-219 to Gly-225, Ala-260 to Ser-268, Thr-276 to Glu-282. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of pre-natal disorders, anomalies and deficiencies. The homology to the hematopoietic lineage switch 2 proteins indicates that The translation product of this gene is useful for the detection and/or treatment of immune system disorders. In addition, the homology to the transferrin receptor indicates that the translation product of the present invention may have utility in the regulation of iron metabolism as well as the numerous genes under the stringent control of physiologic iron levels. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1849 of SEQ ID NO:28, b is an integer of 15 to 1863, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

PRVRSREPVAGAPGCGTAGPPAMATLWGGLLRLGSLLSLSCLALSVLLLAHC QTPPSDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIIIYLSILGLLL LYMVYLTLVEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLARSRSRANV

46

LNKVEYAQQRWKLQVQEQRKSVFDRHVVLS (SEQ ID NO: 225).

Polynucleotides encoding these polypeptides are also provided.

5

10

15

20

25

30

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 72 - 88 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 89 to 167 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

A preferred polypeptide varient of the invention comprise the following amino acid sequence:

MATLWGGLLRLGSLLSLSCLALSVLLLAHCQTPPRISRMSDVNVSALPIKKNS GHIYNKNISQKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIII YLSILGLLLLYMVYLTLVEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLA RSRSRANVLNKVEYGTAALEASSPRAAKSLSLTGMLSSANWGIEFKVTRKKQ ADNWKGTDWVLLGFILIPC (SEQ ID NO: 226). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain tissue, and to a lesser extent in other cell types and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and neurodegenerative diseases of the brain and nervous system, such as depression, schizophrenia, Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, mania, dementia, paranoia, addictive behavior, sleep disorders, epilepsy, transmissible spongiform encephalopathy (TSE), Creutzfeldt-Jakob disease (CJD). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developmental, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic

47

fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 121 as residues: Gln-110 to Pro-120, Val-152 to Val-159. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of developmental, degenerative and behavioral conditions of the brain and nervous system. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of schizophrenia, Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, transmissible spongiform encephalopathy (TSE), Creutzfeldt-Jakob disease (CJD), mania, depression, dementia, paranoia, addictive behavior, obsessive-compulsisve disorder and sleep disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1612 of SEQ ID NO:29, b is an

integer of 15 to 1626, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

The translation product of this gene shares sequence homology with a recently published gene Dysferlin, which is thought to be a skeletal muscle gene that is mutated in Miyoshi myopathy and limb girdle muscular dystrophy (See Genbank Accession No. g3600028, and Nat. Genet. 20 (1), 31-36 (1998)).

This gene is expressed primarily in fetal liver, fetal heart tissue, and T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunodeficiency, tumor necrosis, lymphomas, auto-immunities, cancer, inflammation, anemias (leukemia) and liver disorders, vascular disorders, and cancers (e.g., hepatoblastoma, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, developmental, vascular, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, bile, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g., AIDS), immuno-supressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product is applicable in conditions of general

5

10

15

20

25

30

microbial infection, inflammation or cancer. Expression in liver may suggest a role for this gene product in the treatment and detection of liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Alternatively, the tissue distribution in fetal heart tissue indicates that the protein product of this gene is useful for the diagnosis and treatment of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Additionally, the homology to the dysferlin gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for diseases related to degenerative myopathies that are characterized by the weakness and atrophy of muscles without neural degradation; such as Duchenne and Becker's muscular dystropies. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 591 of SEQ ID NO:30, b is an integer of 15 to 605, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene is expressed primarily in haemopoietic cells and tumor cells, such as pancreatic tumor tissue, and to a lesser extent in bladder cells.

50

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, haemopoietic disorders, diseases of the renal and pancreatic systems, and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemopoietic, pancreatic, and renal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., pancreas, renal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the renal, pancreatic and haemopoietic systems, and cancers thereof. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 917 of SEQ ID NO:31, b is an integer of 15 to 931, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

25

5

10

15

20

51

This gene is expressed primarily in liver tissue, cancer cells and fetal lung tissue, and to a lesser extent in dendritic cells.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic disorders, diseases of developing systems and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetus, metabolic systems and cancers, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, metabolic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 124 as residues: His-44 to Gly-49. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the fetus, metabolic systems and cancers. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

52

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1393 of SEQ ID NO:32, b is an integer of 15 to 1407, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

5

10

15

20

25

30

This gene is expressed primarily in central nervous system tissues and cancers, such as endometrial tumors, and to a lesser extent in other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the CNS and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and cancerous tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 125 as residues: Tyr-16 to Ser-22, Asp-209 to Leu-215. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of diseases of the central nervous system, as well as cancers of tissues where expression of this gene has been observed, such as in endometrial tumors. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1512 of SEQ ID NO:33, b is an integer of 15 to 1526, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

The translation product of this gene shares sequence homology with low-density lipoprotein receptor (See Genbank Accession No. >dbj|BAA24580.1), which is thought to be important in the pathogenesis of atherosclerosis and other disorders.

54

The translation product of this gene also shares sequence homology with a rat homolog of the human CD94 (See Genbank Accession No. gb|AAC10220.1).

This gene is expressed primarily in macrophages, eosinophils, neutrophil and other cells of the haemopoietic and immune system.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune and haemopoietic systems and diseases of the endothelial and vascular system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, haemopoietic and vascular system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 126 as residues: Lys-9 to Ala-17, Met-55 to Leu-61, Tyr-105 to Cys-127, Asp-132 to Lys-141, Ser-165 to Tyr-172, Pro-178 to Met-186, His-222 to Gln-227. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to LDL receptor and rat CD94 homolog indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the immune, haemopoietic and vascular systems. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of hematopoietic disorders. This gene product is primarily expressed in hematopoietic cells and tissues, suggesting that it plays a role in the survival, proliferation, and/or differentiation of hematopoietic lineages. Expression of this gene product in eosinophils and macrophage also strongly indicates a role for this protein in immune

function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1723 of SEQ ID NO:34, b is an integer of 15 to 1737, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 25

5

10

20

25

30

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

MAAAGRLPSSWALFSPLLAGLALLGVGPVPARALHNVTAELFGAEAWGTLA
AFGDLNSDKQTDLFVLRERNDLIVFLADQNAPYFKPKVKVSFKNHSALITSVV
PGDYDGDSQMDVLLTYLPKNYAKSELGAVIFWGQNQTLDPNNMTILNRTFQ
DEPLIMDFNGDLIPDIFGITNESNQPQILLGGNLSWHPALTTTSKMRIPHSHAFI
DLTEDFTADLFLTTLNATTSTFQFEIWENLDGNFSVSTILEKPQNMMVVGQSA
FADFDGDGHMDHLLPGCEDKNCQKSTIYLVRSGMKQWVPVLQDFSNKGTL
WGFVPFVDEQQPTEIPIPITLHIGDYNMDGYPDALVILKNTSGSNQQAFLLENV
PCNNASCEEARRMFKVYWELTDLNQIKDAMVATFFDIYEDGILDIVVLSKGY
TKNDFAIHTLKNNFEADAYFVKVIVLSGLCS NDCPRR (SEQ ID NO: 227).
Polynucleotides encoding these polypeptides are also provided.

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and tissue cell types, through the JAK-STAT

56

signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

5

10

15

20

25

30

This gene is expressed primarily in infant brain and placental tissues, and to a lesser extent in several other tissues including cancers.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, brain disorders and diseases of developing systems and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and fetal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 127 as residues: Leu-56 to Thr-62, Gln-80 to Pro-87, Gly-106 to Gln-113, Pro-122 to Lys-127, Gln-138 to Asn-146. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neural tissues and developing tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the central nervous system, disorders of developing systems, and cancers. The tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful

for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders.

5

10

15

20

25

30

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product is produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product is produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

58

formula of a-b, where a is any integer between 1 to 2228 of SEQ ID NO:35, b is an integer of 15 to 2242, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 26

10

15

20

25

30

Preferred polypeptides of the invention comprise the following amino acid sequence:

MTKREDGGYTFTATPEDFPKKHKAPVIDIGIANTGKFIMTASSDTTVLIWSLK GQVLSTINTNQMNNTHAAVSPCGRFVASCGFTPDVKVWEVCFGKKGEFQEV VRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTXVEYKKKQDPY LLKTGRFEEAAGAXPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNEST RQRLQQQLTQAQETLKSLGALKK (SEQ ID NO: 228). Polynucleotides encoding such polypeptides are also provided.

The gene encoding the disclosed cDNA is thought to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. Recently, another group published this gene, naming it WS beta-transducin repeats protein (See Human Genetics 103 (5), 590-599 (1998); which is hereby incorporated herein by reference), in which it was suggested that the protein is involved in William's Disease.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 12 - 28 of the amino acid sequence referenced in

Table I for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

5

10

15

20

25

30

VIRHEGSTNMELSQMSXLMGLSVLLGLLALMATAAVXRGWLRAGEERSGRP ACQKANGFPPDKSSGSKKQKQYQRIRKEKPQQHNFTHRLLAAALKSHSGNIS CMDFSSNGKYLATCADDRTIRIWSTKDFLQREHRSMRANVELDHATLVRFSP DCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKKHKAPVIDIGIAN TGK

FIMTASSDTTVLIWSLKGQVLSTINTNQMNNTHAAVSPCGRFVASCGFTPDVK VWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWK LWDTXVEYKKKQDPYLLKTGRFEEAAGAXPCRLALSPNAQVLALASGSSIHL YNTRRGEKEECFERVHGECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAM VEEMQGHLKRASNESTRQRLQQQLTQAQETLKSLGALKK (SEQ ID NO: 229). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in testes, synovial sarcoma and fetal tissues, and to a lesser extent in several other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the reproductive and developing systems and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, testicular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, seminal fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

60

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

The tissue distribution in testes tissue, synovial sarcoma, and fetal tissues, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the reproductive and developing systems, and cancers. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

Furthermore, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. This protein is useful for

61

the treatment, detection, and/or prevention of William's Disease. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2221 of SEQ ID NO:36, b is an integer of 15 to 2235, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

5

10

15

20

25

30

Preferred polypeptides of the invention comprise the following amino acid sequence positions 1-363, 2-363, 4-363, 5-363, 30-363, 31-363, 32-363, 75-363, 76-363 and 82-363 of the following amino acid sequence:

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTL

FFAFECRYLAVQLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIE

MEIEATNGAVPQGQRPPPRIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDN

CVERFDHHCPWVGNCVGKRNYRYFYLFILSLSLLTIYVFAFNIVYVALKSLKI

GFLETLKETPGTVLEVLICFFTLWSVVGLTGFHTFLVALNQTTNEDIKGSWTG

KNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSRPPSTQETSSS

LLPQSPAPTEL NSNEMPEDSSTPEEMPPPEPPPEPPQEAAEAEK (SEQ ID NO: 230). Polynucleotides encoding such polypeptides are also provided.

A preferred polypeptide varient of the invention comprises the following amino acid sequence: MLFLFSMATLLRTSFSDPGVIPRALPDEAA

62

FIEMEIEATNGAVPQGQRPPPRIKNFQINNQIVKLKYCYTCKIFRPPRASHCSIC DNCVE RFDHHCPWVGNCVGKRNYRYFYLFILSLSLLTIYVFAFNIVYVALK SLKIGFLETLKGNS WNCSRSPHLLLYTLVRRGTDWISYFPRGSQ PDNQ (SEQ ID NO: 231). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in ovarian and endometrial tumors, fetal liver, spleen and brain tissues, and to a lesser extent in several other tissues and organs.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the developing systems, and cancers of the female reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing, female reproductive and fetal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 129 as residues: Pro-44 to Lys-54, Cys-88 to His-95, Val-103 to Tyr-108, Gln-181 to Ser-190, Thr-192 to Ile-206, Glu-233 to Ser-245, Ser-252 to Ala-286. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in developing systems indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of developing and fetal systems, and cancers. Furthermore, the tissue distribution in ovarian and endometrial tumor tissues indicates that the translation product of this gene is useful for the detection, diagnosis, and/or treatment of cancers of the female reproductive system. Accordingly, preferred are antibodies

63

which specifically bind a portion of The translation product of this gene. Also provided is a kit for detecting these tumors. Such a kit comprises in one embodiment an antibody specific for The translation product of this gene bound to a solid support. Also provided is a method of detecting these tumors in an individual which comprises a step of contacting an antibody specific for The translation product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are more particularly described elsewhere herein.

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2957 of SEQ ID NO:37, b is an integer of 15 to 2971, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

This gene is expressed primarily in normal and cancerous colon tissue, macrophages, endothelial cells and placental tissue, and to a lesser extent in several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, colon cancer and gastrointestinal disorders, immune disorders, vascular diseases and disorders of developing systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

5

10

15

20

25

30

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, vascular and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, gastrointestinal, developmental, vascular, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 130 as residues: Thr-27 to Ser-33. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in macrophage, endothelial and placental tissues, and normal and cancerous colon tissues, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of immune, gastrointestinal and vascular disorders and diseases. Expression of this gene product in colon tissue indicates involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of colon cancer, and cancer in general. Accordingly, preferred are antibodies which specifically bind a portion of the translation product of this gene. Also provided is a kit for detecting colon cancer. Such a kit comprises in one embodiment an antibody specific for The translation product of this gene bound to a solid support. Also provided is a method of detecting colon cancer in an individual The translation which comprises a step of contacting an antibody specific for product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are more particularly described elsewhere herein. Alternatively, the tissue distribution in placental tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders

65

of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product is produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta and endothelial cells also indicates that this gene product is produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Additionally, expression of this gene product in macrophage also strongly indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

66

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1149 of SEQ ID NO:38, b is an integer of 15 to 1163, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

5

10

15

20

25

30

The translation product of this gene shares homology with HNK-sulfotransferase, which directs glycan synthesis (see Genbank Accession no. AF033827).

This gene is expressed primarily in activated T cells, osteoclastoma, and glioblastoma, and to a lesser extent in various other normal and transformed cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation, immune defects, cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hemopoietic systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 131 as residues: Pro-32 to Gly-48, Gln-63 to Thr-69, Pro-77 to Trp-84, Val-88 to Leu-94. Polynucleotides encoding said polypeptides are also provided.

67

The tissue distribution in T-cells and various types of neoplasms indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, study and/or treatment of inflammatory and general immune defects, and various types of neoplasms. Expression of this gene product in T cells strongly indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue distribution in various cancerous tissues indicates that the translation product of the gene is useful for the detection, diagnosis, and/or treatment of these cancers, as well as cancers of other tissues where expression has been observed. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1918 of SEQ ID NO:39, b is an integer of 15 to 1932, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

5

10

15

20

25

30

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

Preferred polypeptides of the invention comprise the following amino acid sequence:

5 LHECLPGSISYLHPRTPWLCLPPQHLSFSTFSPPWQPAMSPVPGTGGPPCGL (SEQ ID NO: 232), and/or MLPLLIICLLPAIEGKNCLRCWPELSALIDYDLQILWVTPGPPTELSQSIHSLFLE DNNFLKPWYLDRDHLEEETAKFFTQVHQAIKTLRDDKTVLLEEIYTHKNLFT ERLNKISDGLKEKGAPPLHECLPGSISYLHPRTPWLCLPPQHLSFSTFSPPWQP 10 AMSPVPGTGGPPCGL (SEQ ID NO: 233). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain, testes and activated T cells, and to a lesser extent in various other normal and transformed cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, reproductive and inflammatory conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neural, immune and male reproductive systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 132 as residues: Gly-41 to Leu-46, Asp-67 to Thr-75, Ile-114 to Ala-123. Polynucleotides encoding said polypeptides are also provided.

69

The tissue distribution in infant brain tissue, testes tissue, and activated T-cells, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis, and/or treatment of neurological, reproductive and immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

10

15

20

25

5

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell type. Alternatively, the tissue distribution in testes tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

30

Furthermore, the tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 867 of SEQ ID NO:40, b is an integer of 15 to 881, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

20

25

30

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

The translation product of this gene shares sequence homology with some human and rodent melanoma and leukocyte specific antigens (see, for example, Genbank accession nos: gi|189384, gi|205898 and gi|180926). In addition, The translation product of this gene shares sequence homology with Tetraspan protein (see, for example, Genbank accession number: GI 3152703). Therefore, it is likely that the polypeptide of this gene shares some biological functions, such as cell-to-cell signaling, adhesion, proliferation, and differentiation with Tetraspan.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 52-68 and 197 - 213 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed

5

10

15

20

25

30

that the protein product of this gene shares structural features to type IIIa membrane proteins.

The transmembrane 4 superfamily (TM4SF) or tetraspan superfamily has at least 16 members (including CD9, CD20, CD37, CD53, CD63, CD81, CD82, A15, CO-029, Sm23, RDS, Uro B, Uro A, SAS, Rom-1, PETA3, and YKK8), is the second biggest subfamily among CD antigen superfamily, and activation antigen of T- cells. All TM4SF member contains four putative transmembrane domains, two extracellular loops, and two short cytoplasmic tails. They are variously expressed on Immature, early, mature, activated lymphocytes, monocytes, macrophages, granulocytes, platelets, eosinophils, basophils, certain leukemic and lymphoma cells, and a variety of other cells and tissues. CD9 cell surface protein is expressed by both hematopoietic and neural cells, and may play a role for CD9 in intercellular signaling in the immune and nervous system. CD63 is a 53-Kd lysosomal membrane glycoprotein that has been identified as a platelet activation molecule, which play important role in cell adhesion of platelets and endothelial cells. Increased mRNA for CD63 antigen was found in atherosclerotic lesions of Watanabe heritable hyperlipidemic rabbits, suggesting a potential role of CD63 in progression of atherosclerosis. CD63 is also a mast cell marker.

CD82 was originally identified as the target of several mAbs inhibitory to syncytium formation induced by human T-cell leukemia virus type I (HTLV-I), the etiological agent of adult T-cell leukemia. Therefore, this gene could be a target for the development of a drug for this leukemia. CD81 is the target of an antiproliferative antibody. A diverse group of human cell lines, including hematolymphoid, neuroectodermal, and mesenchymal cells, express the CD81 protein. Many of the lymphoid cell lines, in particular those derived from large cell lymphomas, were susceptible to the antiproliferative effects of the antibody. CD81 may therefore play an important role in the regulation of lymphoma cell growth. CD9, CD20, CD37, CD63, CD81 and CD82 have been implicated in the regulation of cell growth, adhesion, and signal transduction of B, T lymphocytes and some other non-lymphoid cells. They associate with CD2, CD21, CD4, CD8, MHC Class II molecules, integrins, function as co-receptor for T, B and other lymphoid cells. Some

TM4SF are leukocyte antigens, highly expressed in activated leukocytes, lymphocytes, are highly specific surface marker for lymphoblastic leukemia, lymphoma, melanoma, and neuroblastoma. CD9 has been show to be involved in cell motility and tumor metastasis. These antigen could be a valuable immunogen or target to implement active and passive immunotherapy in patients with cancer. Others have been shown to be involved in inhibition of prostate cancer metastasis. This gene has close homology to C33 antigen (CD82), whic is a member of the transmembrane 4 superfamily (TMSF) and activation antigen of T- cells. C33 Ag (CD82 was originally identified as the target of several mAbs inhibitory to syncytium formation induced by human T-cell leukemia virus type I (HTLV-I), the etiological agent of adult T-cell leukemia. Therefore, this gene could be very important target for developing drug for leukemia. Other members of this family are Sm23, CO-029, R2, TAPA-1, CD9, CD37, CD53, and CD63. CD63 is a 53-Kd lysosomal membrane glycoprotein that has been identified as a platelet activation molecule.

5

10

15

20

25

30

There is strong evidence indicating that CD63 and Pltgp40, a platelet membrane glycoprotein are the same molecule and that CD63/Pltgp40 is identical to the well-characterized, stage-specific melanoma-associated antigen ME491. These antigen could be valuable immunogens or target to implement active and passive immunotherapy in patients with cancer.

This gene is expressed primarily in fetal tissue (kidney, heart, liver, spleen, brain), macrophages, dendritic cells, retina and to a lesser extent in various other tissues, mostly of lymphoid origin or epithelial cell types. In addition This gene is expressed in cancerous tissues (e.g. breast).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders and cancers in a variety of organs and cell types. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at

73

significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developmental, proliferating, immune, hematopoietic, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid, spinal fluid, or amniotic fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 133 as residues: Tyr-123 to Tyr-131, Cys-134 to Ser-145, Tyr-234 to Tyr-244. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution fetal cells and tissues and homology to tumor antigens indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, treatment and diagnosis of lymphoid and epithelial disorders and neoplasms. Additionally, tissue distribution in immune cells and other tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting hematopoesis, including cancers. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity

74

disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

5

10

15

20

25

30

Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue

75

markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1918 of SEQ ID NO:41, b is an integer of 15 to 1932, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

15

20

25

30

10

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

The translation product of this gene shares limited sequence homology with VEGF which is thought to be important in regulation of endothelial cell growth.

Therefore, it is likely that the protein encoded by this gene would share some similar biological functions.

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and tissue cell types, through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, nervous system disease and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 134 as residues: Thr-25 to Pro-46. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

77

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1150 of SEQ ID NO:42, b is an integer of 15 to 1164, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

5

10

15

20

25

30

The translation product of this gene shares sequence homology with human p150 which is thought to be important in signal transduction in neuronal cells. Therefore, it is likely that the protein encoded by this polynucleotide would share some similar biological functions with p150.

This gene is expressed primarily in whole embryo and cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and growth defects/disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS, expression of this gene

78

at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of central nervous system, neurodevelopmental, cognitive, and memory disorders. The tissue distribution also indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked

79

by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

80

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1091 of SEQ ID NO:43, b is an integer of 15 to 1105, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

5

10

15

20

25

30

This gene is expressed primarily in PMA stimulated HL-60 cells and to a lesser extent in 6 week embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting cell differentiation, particularly hematopoietic disorders and/or defects. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 136 as residues: Pro-61 to Asp-68. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study of cellular differentiation and for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia. The tissue distribution also

81

5

10

15

20

25

30

indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine. production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types Aditionally, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue

82

differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1248 of SEQ ID NO:44, b is an integer of 15 to 1262, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in colon.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders and/or defects of the digestive tract including but not limited to cancers of the gastrointestinal tract. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at

5

10

15

20

25

30

significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., gastrointestinal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the digestive system particulary disorders involving the colon. Further, expression of this gene product in colon tissue indicates involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of colon cancer, and cancer in general. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the colon and/or other gastrointestinal tissue including, but not limited to, stomach, small intestine, large intestine, and rectum.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 503 of SEQ ID NO:45, b is an integer of 15 to 517, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

This gene is expressed primarily in blood cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

84

not limited to, immune and hematopoietic diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 138 as residues: Pro-19 to Cys-29, Thr-35 to Glu-44, Val-72 to Lys-78. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis and/or treatment of disorders of the immune and hematopoietic system. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 844 of SEQ ID NO:46, b is an

85

integer of 15 to 858, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

5

10

15

20

25

30

This gene is expressed in multiple tissue systems such as brain, immune cells, prostate, uterus, testes, placenta, and fetal heart as well as in cancerous tissues such as ovarian tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune, reproductive, urogenital, and central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous sytem and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, reproductive, urogenital, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 139 as residues: Tyr-33 to Lys-38. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune, urogenital, reproductive, and central nervous systems. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of diseases of the central nervous system, as well as cancers of tissues where expression of this gene has been observed, such as in ovarian tumors. The tissue

86

distribution in central nervous system tissues also indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo. . Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in

87

5

10

15

20

25

30

proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in uterus indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The protein product is likely involved in preparation of the endometrium of implantation and could be administered either topically or orally. Alternatively, this gene could be transfected in gene-replacement treatments into the cells of the endometrium and the protein products could be produced. Similarly, these treatments could be performed during artificial insemination for the purpose of increasing the likelyhood of implantation and development of a healthy embryo. In both cases this gene or its gene product could be administered at later stages of pregnancy to promote heathy development of the endometrium. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in testes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed

88

against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 6093 of SEQ ID NO:47, b is an integer of 15 to 6107, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

5

10

15

20

25

30

This gene is expressed in a wide range of tissue systems such as brain, immune cells, fetal liver, kidney, testes, breast, and pancreas as well as cancerous tissue such as ovarian tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system, immune system, urogenital, and reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, CNS, urogenital, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

89

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 140 as residues: Met-1 to Ser-7, Asp-32 to Pro-43, Ser-96 to Arg-102. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune, reproductive, urogenital and central nervous systems. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of diseases of the central nervous system, as well as cancers of tissues where expression of this gene has been observed, such as in ovarian tumors. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities,

90

5

10

15

20

25

30

such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue

91

differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.

5

10

15

20

25

30

The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 689 of SEQ ID NO:48, b is an integer of 15 to 703, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

This gene is expressed primarily in macrophages and fetal cells and to a lesser extent in cancerous ovarian tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune diseases, disorders of developing tissues, and cancer.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in

providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of developmental abnormalities and disorders of the immune systems. The tissue distribution cancerous ovaries indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of these tumors. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Expression of this gene product in macrophage cells strongly indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). This gene product may have clinical utility in the treatment of immune dysfunction; in the correction of autoimmunity; in immune modulation; and in the control of inflammation.

The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or

93

other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the

5

10

15

20

25

30

polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution also indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders such as melanomas.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 625 of SEQ ID NO:49, b is an integer of 15 to 639, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

This gene is expressed primarily in neutrophils, bone marrow, brain, and fetal cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disorders, Limbic system disfunction/defects and disorders of the immune system and developing systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, Limbic system and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 142 as residues: Ala-84 to Gln-93. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune, Limbic system, CNS and developing systems. Expression of this gene product in bone marrow, eosinophils, and neutrophils strongly indicates a role for this protein in hematopoiesis and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). This gene product may have clinical utility in the treatment of immune dysfunction; in the correction of autoimmunity; in immune modulation; and in the control of inflammation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune

5

10

15

20

25

30

Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury.

Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Additionally, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the

97

"Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 853 of SEQ ID NO:50, b is an

98

integer of 15 to 867, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

5

10

15

20

25

30

This gene is expressed primarily in ovary and to a lesser extent in fetal tissue, colon, and immune cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, ovarian cancer, gastrointestinal and immune system disorders.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, gastrointestinal, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 143 as residues: Ile-23 to Ala-29. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of ovarian cancer and related metastases. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The tissue distribution in colon tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders involving the gastrointestinal tract. This may include diseases associated with digestion and food absorption, as well as hematopoietic disorders involving the

99

Peyer's patches of the small intestine, or other hematopoietic cells and tissues within the body. Similarly, expression of this gene product in colon tissue indicates again involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of colon cancer, and cancer in general. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their

5

10

20

25

30

interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1555 of SEQ ID NO:51, b is an integer of 15 to 1569, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 42

The translation product of this gene shares sequence homology with retrovirus-related reverse transcriptase pseudogene. In addition, this gene shares homology with human interferon-beta (Genseq accession number T35524; all references available through this accession are hereby incorporated herein by reference), therefore, it is likely that this gene and the protein encoded by this gene shares some similar biological functions with this protein.

This gene is expressed primarily in frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily

101

fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

The tissue distribution in frontal cortex and homology to retrovirus-related reverse transcriptase pseudogene and human interferon-beta indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurodegenerative diseases of the brain, particularly of the frontal cortex. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, multiple schlerosis, cystic fibrosis, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

102

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1182 of SEQ ID NO:52, b is an integer of 15 to 1196, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

5

10

15

20

25

30

This gene is expressed primarily in immune cells, brain, fetal tissue, and cancerous tissues (such as testes, stomach, lung, pancreas, ovaries) and to a lesser extent in other numerous tissues including, but not limited to, testes and kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and immune cells expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 145 as residues: Lys-23 to Lys-35, Met-46 to Tyr-52. Polynucleotides encoding said polypeptides are also provided.

103

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurodegenerative disorders of the frontal cortex, as well as, cancer or a number of tissues including but not limited to testes, stomach, lung, pancreas, and ovaries. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

5

10

15

20

25

30

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell

104

lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent

105

of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 931 of SEQ ID NO:53, b is an integer of 15 to 945, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

106

This gene is expressed primarily in epithelioid sarcoma and to a lesser extent in pancreatic carcinoma, aorta endothelial cells induced with TNF-alpha, and amniotic cells induced with TNF. This gene is also expressed, to a lesser extent, in cancerous lung and ovary tissue and fetal tissue.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, epithelioid sarcoma and related cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., amniotic, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 146 as residues: Tyr-39 to Arg-51. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of certain cancers, including epithelioid sarcoma and pancreatic carcinoma. The tissue distribution in tumors of lung, ovary, and pancreas origins indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of

developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5 .

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types Thus this protein may modulate apoptosis or tissue of degenerative conditions. differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including

108

blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 474 of SEQ ID NO:54, b is an

109

integer of 15 to 488, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

5

10

15

20

25

30

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

PPVPPWISLPLTGSPPRPGFVPVSPFCFSPMTNGHQVLLLLLLTSAVAAGPWPQ VHAGQWGWMCLPPGLPSVQARSGLGGLPGGPQWVPGGARGY (SEQ ID NO: 234). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in fetal and infant tissue, particularly infant brain and fetal liver/spleen libraries, and to a lesser extent in breast, ovary tumor, pharynx carcinoma, endometrial stromal cells, thymus, islet cell tumors, and adult cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and breast, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, developmental, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in developing cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis

110

and treatment of cancer and other proliferative disorders. The expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of

111

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2846 of SEQ ID NO:55, b is an integer of 15 to 2860, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

5

10

15

20

25

30

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

IQQWGDSVLGRRCRDLLLQLYLQRPELRVPVPEVLLHSEGAASSSVCKLDGLI HRFITLLADTSDSRALENRGADASMACRKLAVAHPLLLLRHLPMIAALLHGR THLNFQEFRQQNHLSCFLHVLGLLELLQPHVFRSEHQGALWDCLLSFIRLLLN YRKSSRHLAAFINKFVQFIHKYITYNAPAAISFLQKHADPLHDLSFDNSDLVM LKSLLAGLSLPSRDDRTDRGLDEEGEEESSAGSLPLVSVSLFTPLTAAEMAPY MKRLSRGQTVEDLLEVLSDIDEMSRRRPEILSFFSTNLQRLMSSAEECCRNLA FSLALRSMQNSPSIAAAFLPTFMYCLGSQDFEVVQTALRNLPEYALLCQEHA AVLLHRAFLVGMYGQMDPSAQISEALRILHMEAVM (SEQ ID NO: 235). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in breast cancer, and to a lesser extent in a variety of other cancers, including uterine cancer, synovial sarcoma, and pharynx carcinoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer; proliferative diseases and/or disroders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For

112

a number of disorders of the above tissues or cells, particularly of the breast, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, breast, proliferative, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, breast milk, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 148 as residues: Glu-35 to His-41, Ser-62 to Ala-67, Pro-145 to Leu-155, Glu-157 to Ser-163, Arg-190 to Val-197, Asp-208 to Pro-215, Ser-247 to Pro-252. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in breast cancer tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of cancer. Elevated expression of this gene product in cancers, such as breast cancer, suggest that it is involved in the abnormal proliferation of cells, dedifferentiation, angiogenesis, and other processes that accompany the development of cancer. Thus, therapeutics targeted against this gene product is useful therapeutic products in and of themselves. Alternately, expression of this gene product at elevated levels in breast tissue is reflective of expression within breast lymph nodes, and may suggest a hematopoietic role for this protein. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the

113

polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1545 of SEQ ID NO:56, b is an integer of 15 to 1559, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of this gene shares limited sequence homology with cytochrome-c oxidase. An alternative embodiment is the polypeptide comprising the following amino acid sequence:

MLLKHLQRMVSVPQVKASALKVVTLTANDKTSVSFSSLPGQGVIYNVIVWD PFLNTSAAYIPAHTYACSFEAGEGSCASLGRVSSKVFFTLFALLGFFICFFGHR FWKTELFFIGFIIMGFFFYILITRLTPIKYDVNLILTAVTGSVGGMFLVAVWWR

FGILSICMLCVGLVLGFLISSVTFFTPLGNLKIFHDDGVFWVTFSCIAILIPVVF MGCLRILNILTCGVIGSYSVVLAIDSYWSTSLSYITLNVLKRALNKDFHRAFTN VPFQTNDFIILAVWGMLAVSGITLQIRRERGRPFFPPHPYKLWKQERERRVTNI LDPSYHIPPLRERLYGRLTQIKGLFQKEQPAGERTPLLL (SEQ ID NO: 236).

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

5

10

15

20

25

30

WARLRGPGAHARTSPQPWRGPSPAQAAMGFLQLLVVXVLXSEHRVAGAAE VFGNSSEGLIEFSVGKFRYF

ELNRPFPEEAILHDISSNVTFLIFQIHSQYQNTTVSFSPRRRSPTM (SEQ ID NO: 237). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in keratinocytes, brain, and spinal cord and to a lesser extent in hematopoietic cells and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders; hematopoietic disorders; integumentary disroders; immune dysfunction; learning disabilities. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., integumentary, neural, developmental, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain and spinal cord cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

diagnosis and treatment of a variety of neurological and hematopoietic disorders. For example, elevated levels of expression of this gene product in brain and spinal cord indicates that it is involved in neurodegenerative disorders. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

5

10

15

20

25

30

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Alternately, expression of this gene product in hematopoietic cells indicates that it is involved in the proliferation, differentiation, survival, and activation of all hematopoietic lineages, including stem and progenitor cells. Expression of this gene product in keratinocytes indicates that it is involved in normal skin function, and could be involved in skin disorders, dermatitis, and fibrosis. The protein is useful in detecting, treating, and/or preventing congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's Disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's Disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e.wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox,

5

10

15

20

molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, althletes foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders (i.e., arthritis, trauma, tendonitis, chrondomalacia and inflammation, etc.), autoimmune disorders (i.e., rheumatoid arthritis, lupus, scleroderma, dermatomyositis, etc.), dwarfism, spinal deformation, joint abnormalities, amd chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2050 of SEQ ID NO:57, b is an integer of 15 to 2064, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 48

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

30 PRVRPASPPVRSPARWGSMAGSPLLWGPRAGGVGLLVLLLLGLFRPPPALCA RPVKEPRGLSAASPPLARLALLAASGGQCPEVRRRGRCRPGAGAGASAGAER

QERARAEAQRLRISRRASWRSCCASGAPPATLIRLWAWTTTPTRLQRSSLALC SAPALTLPP (SEQ ID NO: 238). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in human pituitary and to a lesser extent in pineal gland, and other areas of the brain.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pituitary dysfunction; abnormal growth; neurological defects; insufficient milk secretion; abnormal smooth muscle contraction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endocrine, developmental, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, breast milk, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 150 as residues: Pro-36 to Gly-42, Pro-64 to Ala-76, Gly-83 to Ala-90, Ser-100 to Cys-108, Thr-126 to Ser-135. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution primarily in pituitary cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of disorders. Elevated expression of this gene product in the pituitary indicates that it is possibly a hormone-like substance that either controls pituitary development itself, or various processes controlled by the pituitary. These include growth, milk secretion, smooth muscle contraction, diuresis, blood pressure, and homeostasis. Thus, this gene product may have numerous clinical

5

10

15

20

applications. Expression of this gene product in other regions of the brain also indicates that it is involved in normal neurological function, and is useful in the treatment of a variety of neurological disorders. Representative uses are described in the "Biological Activity", "Hyperproliferative Disorders", and "Binding Activity" sections below, in Example 11, 17, 18, 19, 20 and 27, and elsewhere herein. Briefly, the protein can be used for the detection, treatment, and/or prevention of Addison's Disease, Cushing's Syndrome, and disorders and/or cancers of the pancrease (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-,hypoparathyroidism), hypothallamus, and testes. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1036 of SEQ ID NO:58, b is an integer of 15 to 1050, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

PRVRLATPNIWDLSMLFAFISLLVMLPTWWIVSSWLVWGVILFVYLVIRALRL

10

15

20

25

30

WRTAKLQVTLKKYSVHLEDMATNSRAFTNLVRKALRLIQETEVISRGFTLVS AACPFNKAGQHPSQHLIGLRKAVYRTLRANFQAARLATLYMLKNYPLNSES DNVTNYICVVPFKELGLGLSEEQISEEEAHNFTDGFSLPALKVLFQLWVAQSS EFFRRLALLLSTANSPPGPLLTPALLPHRILSDVTQGLPHAHSACLEELKRSYE FYRYFETQHQSVPQCLSKTQQKSRELNNVHTAVRSLQLHLKALLNEVIILEDE LEKLVCTKETQELVSEAYPILEQKLKLIQPHVQASNNCWEEAISQVDKLLRRN TDKKGKPEIACENPHCTVSTFEAAYSTHCRQRSNPRGAGIRSLCR (SEQ ID NO: 239). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 7 - 23 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 24 to 390 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

This gene is expressed primarily in prostate and placenta and to a lesser extent in pancreatic tumors and hematopoietic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancer; pancreatic cancer; prostate dysfunction; hematopoietic disorders; reproductive diseases and/or disorders, and pancreatitis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, prostate, pancease, placental, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, seminal fluid, plasma, urine, synovial fluid and spinal fluid) or

120

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 151 as residues: Pro-85 to Ser-94, Pro-127 to Thr-136, Glu-154 to Glu-160, Phe-240 to Ser-250, Leu-255 to Leu-265, Leu-341 to Lys-351, Thr-372 to Gly-384. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution in prostate and placental cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of reproductive disorders. Elevated expression of this gene product in the prostate indicates that it is involved in normal prostate function, and is a diagnostic marker for prostate cancer. Alternately, expression of this gene product in placenta indicates that it may play a role in normal vascular function, and is involved in such processes as angiogenesis and endothelial cell chemotaxis. Thus, this gene product is useful in the treatment of myocardial infarction, cancer, ischemia, and diabetic retinopathy. Expression of this gene product in placenta may also be indicative of fetal health and development.

Similarly, expression of this gene product in hematopoietic cells indicates that it is involved in the proliferation, differentiation, survival, or activation of all hematopoietic cell lineages. Finally, expression of this gene product in pancreatic cancers indicates that it may play a role in cancer in general, or in pancreatic function. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases

121

and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2519 of SEQ ID NO:59, b is an integer of 15 to 2533, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

When tested against Jurkat and K562 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) and ISRE (interferon-sensitive responsive element) promoter elements, respectively. Thus, it is likely that this gene activates myeloid, leukemia, and to a lesser extent, other immune or hematopoietic cells and tissue cell-types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

122

Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. ISRE is also a promoter element found upstream in many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

5

10

15

20

25

30

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

AAPHPPLLRPLCLWCPLWPAWPLRGRPRSAWKRWPPLPVGPAKLGCSMTTR QPTAVSWPCWLMSSSLSTACLAWTLTGSLAREATRRARSLSPTWNCSARQV PPSPPHSGLGRRGWAHCHLT CLLVTQLFRVGRIHPILSLPLVT (SEQ ID NO: 240). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, vascular diseases; aberrant angiogenesis; neurological disorders; learning disorders; placental insufficiency; and fetal distress. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and neurological systems (CNS/PNS), expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, reproductive, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

123

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 152 as residues: Met-1 to Thr-7, Glu-36 to Ser-43, Pro-46 to Gly-63. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution in brain and placental cells and tissues, combined with the detected GAS and ISRE biological activities, indicates that the protein products of this gene are useful for the diagnosis and/or treatment of a variety of neural, reproductive, and vascular diseases and/or disorders. neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Expression of this gene product in placenta indicates that it may play a role in blood vessel development or function, as the placenta is a highly vascularized organ. Thus, this gene product is involved in such processes as angiogenesis, endothelial cell chemotaxis, and vascular cord formation. Thus, it is useful in the treatment of such conditions as myocardial infarction; ischemia; and cancer. Alternately, expression of this gene product in the brain indicates that it may play a role in the survival, proliferation, or function of neurons, and thus is useful in the diagnosis and treatment of such neurological disorders as ALS, schizophrenia, and

124

Alzheimer's Disease. It may likewise be involved in learning disorders as well. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 885 of SEQ ID NO:60, b is an integer of 15 to 899, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

5

10

15

. 20

25

30

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

LQLASQSAGIKGMSHCARPTFLTLLLASCFWAAAIPNRNVILSVSFRPLHMQ FTLSILVFILRILILLRSFL (SEQ ID NO: 241). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 40 - 56 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 57 to 60 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

125

This gene is expressed primarily in spleen derived from patients with chronic lymphocytic leukemia.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic lymphocytic leukemia; hematopoietic disorders; impaired immune function; cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spleen cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of hematopoietic disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Elevated expression of this protein in the spleens of patients with CLL indicates that it is a useful marker for thi's Disease. Alternately, it is associated with the development and/or progression of the disease, and is a useful target for therapeutic intervention. Additionally, this gene

product may play more general roles in hematopoiesis, and may serve to control cellular decisions regarding proliferation, survival, activation, and/or differentiation of all hematopoietic cell lineages. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1065 of SEQ ID NO:61, b is an integer of 15 to 1079, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 52

5

10

15

25

30

The translation product of this gene shares sequence homology with a putative tyrosine protein kinase from the Chilo iridescent virus. See, for example, Genbank accession no. gi|2738451 (AF003534). Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with tyrosine kinase and signaling proteins. Such activities are known in the art, some of which are described elsewhere herein.

This gene is expressed in a variety of tissues, including microvascular endothelial cells, dendritic cells, and fetal tissues. as well as several tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

127

not limited to, vascular, immune, and developmental diseases and/or disorders, particularly cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., vascular, immune, developmental, proliferative, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 154 as residues: Ala-21 to Lys-31, Arg-41 to Cys-56, Thr-92 to Cys-102, Arg-132 to Val-137, Lys-152 to Ile-159, Pro-199 to Ser-205, Arg-210 to Asp-219, Ser-225 to Lys-230, Tyr-236 to Ala-241, Lys-243 to Leu-249, Thr-375 to Asp-381. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to a tyrosine kinase indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of cancer. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities,

128

. 2

10

15

20

25

30

such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1914 of SEQ ID NO:62, b is an

129

integer of 15 to 1928, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

5

10

15

20

25

30

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 2 - 18 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly cancer and immune suppression. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 155 as residues: Gly-63 to Ser-72. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful as a marker for neutrophil monitoring in cancer and/or immune suppressed patients and/or during chemotherapy or radiation therapy. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and

elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

131

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 767 of SEQ ID NO:63, b is an integer of 15 to 781, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where b is greater than or equal to a + 14.

5

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

This gene is expressed primarily in IL-1 and LPS induced neutrophils, and to a lesser extent, in fetal brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, and neural diseases and/or disorders, particularly cancer and immune suppression. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 156 as residues: Ile-28 to Trp-37, Ser-68 to Lys-81. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful as a marker in neutrophils to monitor patients who are immune suppressed or cancer patients during chemotherapy or radiation therapy. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in

132

regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

133

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1180 of SEQ ID NO:64, b is an integer of 15 to 1194, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

This gene is expressed primarily in prostate.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, urogenital diseases and/or disorders, particularly prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urogenital system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., urogenital, prostate, renal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 157 as residues: Arg-30 to Gln-36. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in prostate cancer cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, treatment and diagnosis of prostate cancer and other urogenital disorders. Moreover, the expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis,

treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1663 of SEQ ID NO:65, b is an integer of 15 to 1677, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where b is greater than or equal to a + 14.

5

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

A preferred polypeptide of the invention comprises the following amino acid sequence:

MVLVLRHPLCARERAFREPGRGLLTRTGQHDGAPAVTAVPGPLGAVAAAEG RRSAWGAGGSSPPRKVLWGDMRGRRAGVDVLGPALSSEAAGAEARGWGM PGMGVGVGASETRGALFLGREGVHGPCPMDGLGPWPWGPW (SEQ ID NO: 242). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in rejected kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders affecting the kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urinary tract, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., urogenital, renal, kidney, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 158 as residues: Ala-30 to Gly-36, Asp-45 to Trp-50, Lys-65 to Cys-71, Pro-80 to Cys-87. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in kidney indicates the protein product of this gene could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. The protein is useful for modulating the immune response to aberrant proteins, as may exist in proliferating cells and tissues. Such modulation of the immune response would also show utility in inhibiting the rejection of transplanted tissues, particularly of the renal system. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1223 of SEQ ID NO:66, b is an integer of 15 to 1237, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

The translation product of this gene shares sequence homology with both human and mouse Fibulin-2 which is an extracellular matrix protein found in heart tissue (See Genbank Accession Nos. emb|CAA57876.1 and emb|CAA53040.1, respectively; all references available through these accessions are hereby incorporated

herein by reference; for example, J. Cell Biol. 123 (5), 1269-1277 (1993)). Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

MGPAVKMWTNAWKGLDDCHYNQLCENTPGGHRCSCPRGYRMQGPSLPCL

DVNECLQLPKACAYQCHNLQGSYRCLCPPGQTLLRDGKACTSLERNGQNVT

TVSHRGPLLPWLRPWASIPGTSYHAWVSLRPGPMALSSVGRAWCPPGFIRQN

GVCTDLDECRVRNLCQHACRNTEGSYQCLCPAGYRLLPSGKNCQDINECEEE

SIECGPGQMCFNTRGSYQCVDTPCPATYRQGPSPGTCFRRCSQDCGTGGPSTL

QYRLLPLPLGVRAHHDVARLTAFSEVGVPANRTELSMLEPDPRSPFALRPLRA

GLGAVYTRRALTRAGLYRLTVRAAAPRHQSVFVLLIAVSPYPY (SEQ ID NO: 243). Polynucleotides encoding these polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

10

15

20

25

30

MRVLVVTIAPIYWALARESGEALNGHSLTGGKFRQSHTWSLLQGAAHDDPV ARGLDPDGLLLLDVVVNGVVPGRAWLTQIFKCRTLKKHYVQTRAWPAVRG LHTALLPGRPPLVPTLQPQHPVQRGPGPPAPAGAAPAGLSYQLGL (SEQ ID NO: 244). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

HASGAFLVVRGEPQGSWGSMTGVINGRKFGVATLNTSVMQEAHSGVSSIHSS IRHVPANVGPLMRVLVVTIAPIYWALARESGEALNGHSLTGGKFRQESHVEF ATGELLTMTQWPGVWIPMASCSSTWWSMALSPDSLADADLQVQDFEEHYV QTGPGQLFVGSTQRFFQGGLPSFLRCNHSIQYNAARGPQPQLVQHLRASAISS AFDPEAEALRFQLATALQAEENEVGCPEGFELDSQGAFCVDVDECAWDAHL CREGQRCVNLLGSYRCLPDCGPGFRVADGAGCEDVDECLEGLDDCHYNQLC ENTPGGHRCSCPRGYRMQGPSLPCLDVNECLQLPKACAYQCHNLQGSYRCL CPPGQTLLRDGKACTSLERNGQNVTTVSHRGPLLPWLRPWASIPGTSYHAWV SLRPGPMALSSVGRAWCPPGFIRQNGVCTDLDECRVRNLCQHACRNTEGSY QCLCPAGYRLLPSGKNCQDINECEEESIECGPGQMCFNTRGSYQCVDTPCPAT YRQGPSPGTCFRRCSQDCGTGGPSTLQYRLLPLPLGVRAHHDVARLTAFSEV

138

GVPANRTELSMLEPDPRSPFALRPLRAGLGAVYTRRALTRAGLYRLTVRAAA PRHQSVFVLLIAVSPYPY (SEQ ID NO: 245). Polynucleotides encoding these polypeptides are also provided.

When tested against U937 and Jurkat cell lines, supernatants removed from cells containing this gene repeatedly activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid, T-cells, and to a lesser extent, other immune and hematopoietic cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in kidney.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders affecting the kidney and renal system.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urinary tract, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., renal, urogenital, kidney, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 159 as residues: Lys-32 to Ser-37, His-89 to Gly-94, Asn-124 to Gln-130, Ala-163 to Val-168, Cys-196 to Arg-201, Gln-244 to Gln-264,

139

His-288 to Tyr-294, Leu-314 to Gln-319, Ala-392 to Ser-399, Pro-412 to Asp-419, Ala-452 to Pro-460, Arg-466 to Thr-473. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in rejected kidney, the homology to the conserved Fibulin-2 protein, in addition to the detected GAS biological activity, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting kidneys, particularly proliferative disorders. Representative uses are described here and elsewhere herein. The protein product of this gene could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1920 of SEQ ID NO:67, b is an integer of 15 to 1934, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where b is greater than or equal to a + 14.

25

5

10

15

20

10

20

25

30

Preferred polypeptides of the invention comprise the following amino acid sequence:

MGEKFLLLAMKENHPECFCKILKILHCMDPGEWLPQTEHCVHLTPKEFLIWT MDIASNERSEIQSVALRLASKVISHHMQTCVENRELIAAELKQWVQLVILSCE DHLPTESRLAVVEVLTSTTPLFLTNPHPILELQDTLALWKCVLTLLQSEEQAV RDAATETVTTAMSQENTCQSTEFAFCQVDASIALALALAVLCDLLQQWDQL APGLPILLGWLLGESDDLVACVESMHQVEEDYLFEKAEVNFWAETLIFVKYL CKHLFCLLSKSGWRPPSPEMLCHLQRMVSEQCHLLSQFFRELPPAAEFVKTV EFTRLRIQEERTLACLRLLAFLEGKEGEDTLVLSVWDSYAESRQLTLPRTEAA C (SEQ ID NO: 246). Polynucleotides encoding such polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MGEPNRHPSM
FLLLLVLERLYASPMDGTSSALSMGPFVPFIMRCGHSPVYHSREMAARALVP
FVMIDHIPNTIRTLLSTL

PSCTDQCFRAKPHSWGHFSRFFHLLQAYSDSKTRNEFRLPARAD (SEQ ID NO: 247). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

following amino acid sequence:

MTGREFFSRFPELYPFLLKQLETVANTVDSDMGEPNRHPSMFLLLLVLERLY
ASPMDGTSSALSMGPFVPFIMRCGHSPVYHSREMAARALVPFVMIDHIPNTIR

TLLSTLPSCTDQCFRQNHIHGTLLQVFHLLQAYSDSKHGTNSDFQHELTDITV

CTKAKLWLAKRQNPCLVTRAVYIDILFLLTCCLNRSAKDNQPVLESLGFWEE

VRGIISGSELITGFPWAFKVPGLPQYLQSLTRLAIAAVWAAAAKSGERETNVPI

SFSQLLESAFPEVRSLTLEALLEKFLAAASGLGEKGVPPLLCNMGEKFLLLAM

KENHPECFCKILKILHCMDPGEWLPQTEHCVHLTPKEFLIWTMDIASNERSEIQ

SVALRLASKVISHHMQTCVENRELIAAELKQWVQLVILSCEDHLPTESRLAVV

EVLTSTTPLFLTNPHPILELQDTLALWKCVLTLLQSEEQAVRDAATETVTTAM

SQENTCQSTEFAFCQVDASIALALALAVLCDLLQQWDQLAPGLPILLGWLLG

ESDDLVACVESMHQVEEDYLFEKAEVNFWAETLIFVKYLCKHLFCLLSKSG

5

10

15

20

25

30

WRPPSPEMLCHLQRMVSEQCHLLSQFFRELPPAAEFVKTVEFTRLRIQEERTL ACLRLLAFLEGKEGEDTLVLSVWDSYAESRQLTLPRTEAAC (SEQ ID NO: 248). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 144 - 160, and 462 - 478 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIa membrane proteins. Included in this invention as a preferred domain is the formate and nitrite transporters domain, which was identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). A number of bacterial and archaebacterial proteins involved in transporting formate or nitrite have been shown [1] to be related: - focA and focB, from Escherichia coli, transporters involved in the bidirectional transport of formate. - fdhC, from Methanobacterium formicicum and thermoformicicum, a probable formate transporter. - nirC, from Escherichia coli and Salmonella typhimurium, a probable nitrite transporter. - Bacillus subtilis hypothetical protein yrhG. - Bacillus subtilis hypothetical protein ywcJ (ipa-48R). These transporters are proteins of about 280 residues and seem to contain six transmembrane regions. As signature patterns, we selected two conserved regions. The first one is located in what seems to be a cytoplasmic loop between the second and third transmembrane domains; the second is part of the fourth transmembrane region. The 70 Kd yeast hypothetical protein YHL008c is highly similar, in its Nterminal section, to the prokaryotic members of this family. The concensus pattern is as follows: [LIVMA]-[LIVMY]-x-G-[GSTA]-[DES]-L-[FI]-[TN]-[GS].

Preferred polypeptides of the invention comprise the following amino acid sequence: IISGSELITG (SEQ ID NO: 249). Polynucleotides encoding these polypeptides are also provided. Further preferred are polypeptides comprising the formate and nitrite transporter domain of the sequence referenced in Table for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C- terminal to the formate and nitrite transporter domain. Alternatively, the additional contiguous amino acid residues is both N-terminal and C-terminal to

the formate and nitrite transporter domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to formate and nitrite transporter proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with formate and nitrite transporter proteins. Such activities are known in the art, some of which are described elsewhere herein. It is believed that this gene maps to chromosome 2. Accordingly, polynucleotides derived from this gene are useful in linkage analysis as markers for chromosome 2.

5

10

15

20

25

30

This gene is expressed primarily in cells of the immune system, primarily T-cells and to a lesser extent in spleen, liver, thymus, tonsils, and testis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly disorders affecting hematopoesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of hematopoetic cells, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 160 as residues: Gly-2 to Pro-8, Ser-82 to His-92, Tyr-107 to Asp-117, Arg-162 to Pro-169, Ser-224 to Thr-229, Leu-310 to His-315, Ser-333 to Glu-338, Glu-381 to Ser-388, Gln-428 to Ala-433, Met-446 to Thr-455, Ser-548 to Ser-554, Gly-613 to Asp-618, Ser-627 to Gln-633. Polynucleotides encoding said polypeptides are also provided.

143

The tissue distribution in immune cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting hematopoesis, including cancers. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

15

25

30

related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3286 of SEQ ID NO:68, b is an integer of 15 to 3300, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 59 10

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

VDGIDKLDIEFLQQFLETHSRGPRLHSPGHASQEATPGANMSSGTELLWPGAA LLVLLGVAASLCVRCSRPGAKRSEKIYQQRSLREDQQSFTGSRTYSLVGQAW PGPLADMAPTRKDKLLQFYPSLEDPASSRYQNFSKGSRHGSEEAYIDPIAMEY YNWGRFSKPPEDDDANSYENVLICKQKTTETGAQQEGIGGLCRGDLSLSLAL KTGPTSGLCPSASPEEDEGI (SEQ ID NO: 250). Polynucleotides encoding these polypeptides are also provided. 20

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 10 - 26 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

This gene is expressed primarily in bone marrow, CD34 positive cells, and immune cells, including, neutrophils, T-cells, B-cells, macrophages, monocytes, and dendritic cells and to a lesser extent in brain and tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting the immune and hematopoietic systems, particularly hematopoesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the the immune system and hematopoeitic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 161 as residues: Ser-29 to Thr-57, Pro-74 to Lys-79, Pro-85 to Glu-107, Tyr-118 to Tyr-136, Gln-144 to Gln-152, Ala-182 to Glu-188. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune and hematopoietic cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting the immune system and hematopoesis. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for

immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities.

5

10

15

20

25

30

Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors);

10

15

20

hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures.

Based upon the the proteins immune cell specific message distribution, it may be involved

in many aspects of the immune response, especially its initial stages, inflammation, allograft rejection, infectious disease response etc. The expression of this clone is frequently found in the hematopoietic cell cDNA libraries. Thus, this factor could be involved in the control of hematopoietic cell proliferation, differentiation, and function. Based on this one can postulate its use in the management of anemias, leukemias, neutropenia, thrombocytopenia, autoimmune diseases, blood tissue engraftment, and poikilothromerythromatosis. Furthermore, the

protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1783 of SEQ ID NO:69, b is an integer of 15 to 1797, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where b is greater than or equal to a + 14.

25

15

20

25

30

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

VLWREASALVLSNRLSSGLLHDLLLQPAIHSRLFPRRSRGLSEGEGSSVSLQRS RVLSAMKHVLNLYLLGVVLTLLSIFVRVMESLEGLLESPSPGTSWTTRSQLAN TEPTKGLPDHPSRSM (SEQ ID NO: 251). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in immune cells including activated T cells, macrophages, jurkat cells, bone marrow cells, and osteoblasts and to a lesser extent in kidney cortex, brain, placenta and lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly inflammation and diseases related to inflammatory activity. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 162 as residues: Pro-34 to Met-63. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating or diagnosing disease related to the normal or abnormal activation of T cells.

Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1359 of SEQ ID NO:70, b is an integer of 15 to 1373, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

5

10

15

20

25

30

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

YTFHTQIFLDFPMIFLTVLPLAFLFLHSGFYHYISFSCLFSLSLALFFFLDVATFR RPGQLFCERSVLFDMFHFGFVSLFLHEWIQAKHFWAGLF IVLPSDVFFSVHHLEAPDGSFPNIAKLSLIILLR (SEQ ID NO: 252).

Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 2 - 18 and 22 - 38 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIa membrane proteins.

This gene is expressed in many tissues including brain, liver, prostate, testes, cartilage, gall bladder. Expression is also seen in a number of tumors including colon carcinoma, pancreas tumor, osteoclastoma, ovarian cancer, B cell lymphoma and acute lymphocytic leukemias.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of various organs including the pancreas, colon, and bone. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the

major organs, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hepatic, metabolic, reproductive, testicular, skeletal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tumors and proliferative tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating or diagnosing tumors of several major organs including the pancreas and large intestine. This protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the

protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1565 of SEQ ID NO:71, b is an integer of 15 to 1579, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

5

10

15

20

25

30

This gene is expressed primarily in dendritic cells and fetal liver/spleen and to a lesser extent in many tissues including tonsils, fetal lung, stromal cell lines, bone marrow cell lines, placenta and tumors including hepatocellular carcinoma, pancreas tumor and osteosarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of the immune and hematopoietic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,

153

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

The tissue distribution in dendritic cells and fetal liver/spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnaosing and treating disorders of the immune system particularlly related to the control and generation of precursor cells. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

154

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1014 of SEQ ID NO:72, b is an integer of 15 to 1028, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

5

10

15

20

25

30

This gene is expressed primarily in adrenal gland tumor and endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine and vascular diseases and/or disorders, particularly diseases associated with the vascular endothelium. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endocrine, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in endothelial cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and treating disorders that involve the vascular system including diseases such as atherschlerosis, neoangiogenesis associated with tumor growth and conditions associated with inflammation. Moreover, the protein is useful in the detection, treatment, and/or prevention of a variety of vascular disorders and conditions, which include, but are not limited to miscrovascular disease, vascular leak syndrome, aneurysm, stroke, embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis. Alternatively, the protein is useful in the treatment, detection, and/or

10

15

20

30

prevention of metabolic disorders, particularly lethargy and depression. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3660 of SEQ ID NO:73, b is an integer of 15 to 3674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

The translation product of this gene is related to bovine PAM precursor. See Genbank record gi|163482 incorporated herein by reference. Moreover, see following patent publications are also incorporated herein by reference: J04311386 and W08902460. Many bioactive peptides terminate with an amino acid alpha-amide at their COOH terminus. The enzyme responsible for this essential posttranslational modification is known as peptidyl-glycine alpha-amidating monooxygenase or PAM. An NH2-terminal signal sequence and short propeptide precede the NH2 terminus of purified PAM. The sequences of several PAM cyanogen bromide peptides were localized in the NH2-terminal half of the predicted protein. The forms of PAM purified from bovine neurointermediate pituitary is generated by endoproteolytic cleavage at a subset of the 10 pairs of basic amino acids in the precursor. High levels of PAM mRNA have been found in bovine pituitary and cerebral cortex. In

156

corticotropic tumor cells, levels of PAM mRNA and pro-ACTH/endorphin mRNA are known to be regulated in parallel by glucocorticoids and CRF.

This gene is expressed primarily in endometrial tumors, dendritic cells, a multiple sclerosis library, kidney, hematopoietic cells, melanocytes, osteoblasts, the spleen, colon, ovary, stromal cells, fetal and adult brain, heart, and in tissues undergoing wound repair.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometriosis, endometrial cancer, multiple sclerosis, hematopoietic diseases, bone disease, and wound healing. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly the hematopoietic system and female reproduction. expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, immune, hematopoieticm integumentary, skeletal, gastrointestinal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in dendritic and hematopoietic cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapuetic or diagnostic agent i's Diseases of hematopoietic origin as well as the female reproductive track due to the gene's primary pattern of expression. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses

include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

5

10

15

20

25

30

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The protein may also have a very wide range of biological activities. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 2783 of SEQ ID NO:74, b is an integer of 15 to 2797, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 65

10

20

25

30

The translation product of this gene shares sequence similarity with several G-protein coupled receptors (See Genbank Accession No. gb|AAC77910.1| (AF061443); all references available through this accession are hereby incorporated herein by reference; for example, Mol. Endocrinol. 12, 1830-1845 (1998)). G-protein coupled receptors are well known in the are and affect a variety of functions. In particular, the translation product of this gene shares similarity with Follical Stimulating Hormone Receptor.

Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

15 GTRFPTGETPSLGFTVTLVLLNSLAFLLMAVIYTKLYCNLEKEDLSENSQSSMI KHVAWLIFTNCIFFCPVAFFSFAPLITAISISPEIMKSVTLIFFP (SEQ ID NO: 253). Polynucleotides encoding such polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MIKHVAWLIFTNCIFFCP

VAFFSFAPLITAISISPEIMKSVTLIFFPCLLA (SEQ ID NO: 254). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

GTRFPTGETPSLGFTVTLVLLNSLAFLLMAVIYTKLYCNLEKEDLSENSQSSMI KHVAWLIFTNCIFFCPVAFFSFAPLITAISISPEIMKSVTLIFFPLPACLNPVLYVF FNPKFKEDWKLLKRRVTKKSGSVSVSISSQGGCLEQDFYYDCGMYSHLQGN LTVCDCCESFLLTKPVSCKHLIKSHSCPALAVASCQRPEGYWSDCGTQSAHS DYADEEDSFVSDSSDQVQACGRACFYQSRGFPLVRYAYNLPRVKD (SEQ ID

NO: 255). Polynucleotides encoding these polypeptides are also provided.

10

15

30

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 43 - 59 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 60 to 207 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins. Included in this invention as preferred domains are Zinc finger, C2H2 type domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). 'Zinc finger' domains [1-5] are nucleic acid-binding protein structures first identified in the Xenopus transcription factor TFIIIA. These domains have since been found in numerous nucleic acid-binding proteins. A zinc finger domain is composed of 25 to 30 amino-acid residues. There are two cysteine or histidine residues at both extremities of the domain, which are involved in the tetrahedral coordination of a zinc atom. It has been proposed that such a domain interacts with about five nucleotides. A schematic representation of a zinc finger domain is shown below:

 \mathbf{x} \mathbf{x} Х Х 20 Х X Х Х X X \mathbf{C} Н $x \setminus / x$ 25 Zn x x / x \mathbf{C} H $x \times x \times x$ XXXXX

Many classes of zinc fingers are characterized according to the number and positions of the histidine and cysteine residues involved in the zinc atom

10

30

coordination. In the first class to be characterized, called C2H2, the first pair of zinc coordinating residues are cysteines, while the second pair are histidines. A number of experimental reports have demonstrated the zinc- dependent DNA or RNA binding property of some members of this class. Some of the proteins known to include C2H2-type zinc fingers are listed below. We have indicated, between brackets, the number of zinc finger regions found in each of these proteins; a '+' symbol indicates that only partial sequence data is available and that additional finger domains is present. In addition to the conserved zinc ligand residues it has been shown that a number of other positions are also important for the structural integrity of the C2H2 zinc fingers. The best conserved position is found four residues after the second cysteine; it is generally an aromatic or aliphatic residue. The concensus pattern is as follows: C-x(2,4)-C-x(3)-[LIVMFYWC]-x(8)-H-x(3,5)-H.

Preferred polypeptides of the invention comprise the following amino acid sequence: CDCCESFLLTKPVSCKHLIKSH (SEQ ID NO: 256). Polynucleotides encoding these polypeptides are also provided. Further preferred are polypeptides 15 comprising the Zinc finger, C2H2 type domain of the sequence referenced in Table for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C- terminal to the Zinc finger, C2H2 type domain. Alternatively, the additional contiguous amino acid residues is both N-terminal and 20 C-terminal to the Zinc finger, C2H2 type domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to zinc finger proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with G-coupled proteins, their receptors, and 25 zinc finger proteins. Such activities are known in the art, some of which are described elsewhere herein.

This gene is expressed primarily in adult and fetal liver, human placenta, colon carcinoma cell lines and fibroblasts and to a lesser extent in the fetal and adult brain, the developing nervous system, lung, pancreas, salivary gland, breast tissue, and dendritic cells.

161

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the liver, developmental abnormalities, neurologic diseases, lung cancer, pancreatic cancer, and colon cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological and hepatic origin, as well as the proliferation and/or differentiation of numerous types of tissues. expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, immune, hematopoietic, neural, gastrointestinal, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 167 as residues: Pro-62 to Asp-67, Arg-74 to Gly-80, Gln-146 to Glu-168. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in fetal liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for a diagnositic marker or therapeutic in a wide variety of disease states, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

162

5

10

15

20

25

30

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the protein expression in placental and brain tissue indicates the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. The protein is useful in the treatment, detection, and/or prevention of bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's Disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, severe mental retardation and dyskinesias, such as Huntington's Disease or Gilles de la Tourette's syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:75 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

163

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2689 of SEQ ID NO:75, b is an integer of 15 to 2703, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:75, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

5

10

15

20

25

30

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

ALENSGSPGLQDSARAHFNXSLRSFSFLRNQMYIFELSLYLEGTSFVVVLLFLL ISVSLDSPPTTKGWDSVLHIWVPLIVQ (SEQ ID NO: 257). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in placenta and in hematopoietic cells, especially those of T-cell and monocyte origin and to a lesser extent in the brain, endothelial cells, and the lungs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopietic, vascular, and developmental diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., vascular, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 168 as residues: Ser-30 to Trp-37. Polynucleotides encoding said polypeptides are also provided.

5

10

15

20

25

30

The tissue distribution in hematopoietic cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for therapeutic and/or diagnostic intervention in hematopoietic and developmental disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of

165

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 728 of SEQ ID NO:76, b is an integer of 15 to 742, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

5

10

15

20

25

30

This gene is expressed primarily in the prostate and to a lesser extent in in human B-cell lymphomas.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancer and diseases of hematopoietic origin, particularly of B-cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., prostate, reproductive, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, seminal fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 169 as residues: Asp-33 to Lys-42. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in prostate tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapeutic or diagnostic marker for prostate cancer and disorders involving hematopoietic cells, especially

166

those of B-cell origin. Moreover, the expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

5

10

15

20

25

30

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. The protein is useful in modulating the immune response to aberrant proteins and polypeptides, as may exist in rapidly proliferating cells and tissues. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1811 of SEQ ID NO:77, b is an integer of 15 to 1825, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

GHESICGSCRSWIYFSIRCRRRMRPWWSLLLEACATCAQTGPTRSTSCTQEVS HSSSTAYPAPMRRRCCL PSPRSCT (SEQ ID NO: 258). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

This gene is expressed primarily in the brain and the developing embryo and to a lesser extent in the heart, colon, adipose tissue, kidney, mammary tissue, activated T-cells and dendritic cells.

5

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological diseases, developmental conditions, colon cancer, and hematopoietic diseases, especially of T-cell origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developmental, cardiovascular, adipose, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 170 as residues: Thr-18 to Cys-26, Glu-29 to Thr-36, Ser-50 to Thr-55. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain, combined with the detected GAS biological activity, indicates that polynucleotides and polypeptides corresponding to this gene are useful for therapeutic and/or diagnostic agents in neurological diseases, developmental abnormalities, colon cancer, and hematopoietic diseases, especially those of T-cell origin. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal

169

cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

5

10

15

20

25

30

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1660 of SEQ ID NO:78, b is an integer of 15 to 1674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 2 - 18 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type II membrane proteins.

170

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

5 KRAGVEVGGLVMALAGSVFVLGGVLVLCVERNGEGEMGWPQHLPKSQPLS PPVAVRRCSFERSWIDLLVETSSSMVTCRQQVGTPNGMEGRGGGPKTTFPIRL QLSGACAVRPEIQWEV (SEQ ID NO: 259). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in activated monocytes, dendritic cells, and in the tonsils.

10

15

20

25

30

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly leukemia, lymphomas, tumors of hematopoietic origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 171 as residues: Gln-30 to Leu-38, Asn-75 to Thr-86. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in activated monocytes, dendritic cells, and tonsils indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapeutic and/or diagnostic agent for leukemias, lymphomas, and other diseases associated with cells of hematopoietic origin. Representative uses are described in the

"Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

172

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2177 of SEQ ID NO:79, b is an integer of 15 to 2191, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

5

10

15

20

25

30

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

This gene is expressed primarily in the placenta, brain, and liver and to a lesser extent in most other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic, neurological, vascular, and developmental diseases and/or disorders, particularly cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or

cell types (e.g., hematopoietic, neurological, vascular, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, bile, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful therapeutic and/or diagnostic agent in a multitude of disease states, particularly those involving the immune and neurologic systems. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the protein is useful in the detection, treatment, and/or prevention of a variety of vascular disorders and conditions, which include, but are not limited to miscrovascular disease, vascular leak syndrome, aneurysm, stroke, embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis.

Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement.

10

15

20

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1321 of SEQ ID NO:80, b is an integer of 15 to 1335, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 71

The translation product of this gene shares sequence homology with the murine Fig1 (interleukin-four induced gene 1) which shares homology to the monoamine oxidases, particularly in domains responsible for FAD binding. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: QDWKAERSQDPFEKCMQDPDYEQLLKVTILEADNRIGGRIFTYRDQXTGWIG ELGAMRMPSSHRILHKLCQGLGLNLTKFTQYDKNTWTEVHEXKLRNYVVEK VPEKLGYALRPQEKGHSPEDIYQMALNQALKDLKALGCRKAMKKFERHTLL EYLLGEGNLSRPAVQLLGDVMSEDGFFYLSFAEALRAXSCLSDRLQYSRIVG GWDLLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNLKVLKAD VVLLTASGPAVKRITFS (SEQ ID NO: 260), and/or

25 LPRHMQEALRRLHYVPATKVFLSFRRPFWREEHIEGGHSNTDRPSRMIFYPPP
REGALLLASYTWSDAAAAFAGLSREEALRLALDDVAALHGPVVRQLWDGT
GVVKRWAEDQHSQGGFVVQXPALWQTEKDDWTVPYGRIYFAGEHTAYPHG
WVETAVKSALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSH
DLAKEEGS (SEQ ID NO: 261). Polynucleotides encoding such polypeptides are

30 also provided.

175

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

MAPLALHILLVLVPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVTIL
EADNRIGGRIFTYRDQXTGWIGELGAMRMPSSHRILHKLCQGLGLNLTKFTQ
YDKNTWTEVHEXKLRNYVVEKVPEKLGYALRPQEKGHSPEDIYQMALNQA
LKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSEDGFFYL
SFAEALRAXSCLSDRLQYSRIVGGWDLLPRALLSSLSGLVLLNAPVVAMTQG
PHDVH

5

10

15

20

25

30

VQIETSPPARNLKVLKADVVLLTASGPAVKRITFSPRCPATCRRRCGGCTTCR PPRCS (SEQ ID NO: 262). Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with monoamine oxidases, disintegrins, metalloproteinases, and apoptosis modulating proteins. Such activities are known in the art, some of which are described elsewhere herein. Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 235 - 251 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 252 to 319 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

This gene is expressed primarily in hematopoietic cells, particularly in dendritic cells, and activated monocytes and to a lesser extent in T-cells, endothelial cells, and cells associated with ulcerative colitis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, leukemias, lymphomas, and diseases associated with antigen presenting cells, in addition to apoptosis dependant events. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression

of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 173 as residues: Gln-22 to Gln-44, Ala-90 to Gly-95, Lys-137 to Trp-146, Arg-171 to Asp-181, Glu-370 to Ser-380, Asp-447 to Gly-452, Gln-463 to Trp-469, Asn-504 to Ala-510, Asp-512 to His-519, Ala-541 to Val-550, Asn-558 to His-566. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution immune and hematopoietic cells and tissues, combined with the homology to the murine Fig 1 gene indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapeutic and/or diagnostic agent for hematopoietic diseases, especially those associated with antigen presenting cells. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatou's Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's

Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1853 of SEQ ID NO:81, b is an integer of 15 to 1867, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:81, and where b is greater than or equal to a + 14.

		Last	AA	o	ORF	106		98		302		224		55		189		123		99		128	
		First	AA of	Secreted	Portion	27		58		38		38		21		23		2		39		27	
	Last	AA	of	Sig	Pep	26		57		37		37		20		22		_		38		26	
	First Last	AA	Jo	Sig	Pep	-		-								-		_		-		_	
	AA	SEQ	А	Ö	Y	103		104		105		174		901		107		175		108		109	
S' NT	of	First SEQ	AA of	Start Signal NO:	Pep	232		69		89		74		312		251		448		62		290	
		S' NT	oę	Start	Codon	232		69		89		74		312		251		448		62		290	
	3' NT	of	Clone	Seq.		1113		983		973		984		1458		2005		2664		933		1480	
	5' NT 3' NT	of	Clone Clone	Seq.		1		102		-		_		-		_			_	I		288	
			Total	Ę	Seq.	1113		983		973		984		1458		2005		2664		943		1503	
	Ł	SEQ	白	Ö.	×	11		12		13		82		14		15		83		91		11	
					Vector	pSport1		Lambda ZAP	П	Uni-ZAP XR		Uni-ZAP XR		pCMVSport	3.0	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR	
		ATCC	Deposit	Nr and	Date	209878	86/81/50	209878	05/18/98	209878	05/18/98	209878	05/18/98	209878	05/18/98	209878	05/18/98	828602	05/18/98	209878	05/18/98	209878	05/18/98
				cDNA	Clone ID	HISCN02		HHGDM70		HHPGO40		HHPGO40		HAMGG68		HAPOM49		HAPOM49		HBGBA69		HBJFJ26	
				Gene	Š.			2		3		3		4		5		5		9		7	

LN	L	ZZ			5.NT 3'NT	3' NT		5' NT of	AA	First Last	Last		
	ATCC		SEQ		Jo		5' NT	First SEQ		AA	AA	First	Last
	Deposit			Total	Total Clone Clone	Clone	of	AA of ID		Jo	oę	AA of	AA
cDNA	Nr and		NO:	K	Seq.	Seq.	Start	Start Signal NO:		Sig	Sig	Secreted	of
Clone ID	Date	Vector	X	Seq.		$\overline{}$	Codon	Pep	Y	Рер	Pep	Portion	ORF
HBJFJ26	209878	Uni-ZAP XR	84	1328	413	1305	169	165	176	-	20	21	59
	05/18/98												
нсерн38	209878	Uni-ZAP XR	81	1512	1	1438	222	222	110	_	26	27	89
	05/18/98												
HDPOJ08	209878	pCMVSport	61	1655		1655	159	159	111	-	81	61	122
	05/18/98	3.0											
HDPRX82	209878	pCMVSport	20	2525	1	2525	128	128	112	_	32	33	82
	05/18/98	3.0											
HELGK31	209878	Uni-ZAP XR	21	1396	25	1334	500	209	113		53	30	344
	05/18/98												
HCNUA40	86826	pBluescript	88	1342	949	1237	096	096	177	_	33	34	105
	02/26/97												
	209044												
	05/15/97												
HFPCX64	209878	Uni-ZAP XR	22	1069	-	1069	181	181	114	_	28	53	181
	05/18/98												
HFPCX64	209878	209878 Uni-ZAP XR	98	1154	84	1154	257	257	178	_	28	29	87
	05/18/98												
HCEBW71	209225	Uni-ZAP XR	87	1197	141	1197	257	257	179	-	28	29	87
	08/28/97												

									5' NT					
				Ä		5' NT 3' NT	3' NT		of	AA	First Last	Last		
		ATCC		SEQ		of	of	5' NT	First SEQ	SEQ	AA	AA	First	Last
		Deposit		A		Total Clone Clone	Clone		AA of D	А	oę	o	-	AA
Gene	cDNA	Nr and		Ö.	Ę	Seq.	Seq.		Start Signal NO:	ÖN			Secreted	Jo
No.	Clone ID	Date	Vector	×	Seq.			Codon	Pep	Y	Pep	Pep	Portion	ORF
13	HFXD060	209878	Lambda ZAP	23	1658	1	1658	131	131	115	-	46	47	115
		05/18/98	П											
14	HHEPG41	209878	pCMVSport	24	1077	385	1043	514	514	911	-	35	36	20
		05/18/98	3.0											
14	HAUAI83	209626	Uni-ZAP XR	88	910	-	988	253	253	180	-	37	38	49
		02/17/98												
14	HJPAZ83	209626	Uni-ZAP XR	68	9/01	398	1076		575	181		11	12	23
		02/17/98												
15	HKGAH42	209878	pSport1	25	1205	_	1205	143	143	117		21	22	63
		05/18/98												
91	HMIAP86	209878	Uni-ZAP XR	56	1674	13	1674	182	182	118	_	19	20	334
		05/18/98												
17	HMUAP70	209878	pCMVSport	27	1965	531	1914	183	183	119	_	91	11	221
		05/18/98	3.0											
17	HMUAP70	209878	pCMVSport	96	1842	407	1783	413	413	182	-	25	56	103
		05/18/98	3.0											
17	HAGFY16	97923	Uni-ZAP XR	16	1963	209	1922	251	251	183	-	28	59	198
		03/07/97								·				
•		209071												
		05/22/97												

	Last	AA	of	ORF	70		82		45		472		33		167		231				46		108	
	First	AA of	Secreted	Portion ORF	45		26		27		25				31		31				33		22	
I act	AA	oę	Sig	Pep	44		25		26		24				30		30			_	32		21	
Firet I act	AA	of	Sig	Pep	1		-		1		1		1		1		ī				-		1	
V	SEQ	А	NO:	Y	184		185		186		120		187		121		881				122		123	
5. NT	First SEQ	AA of	Start Signal NO:	Pep	170		413		128		66		289		89		129				172		46	
	5' NT	oę	Start	Codon	170		413		128		66		687		89		129				172		46	
3, NT	of	Clone	Seq.		1487		1637		1786		1863		1132		1626		1772				909		931	
5, NT 3, NT	of	Total Clone Clone	Seq.		62		394		87		∞		472		I		69			:	1		329	
		Total	Ĭ	Seq.	1487		1653		1830		1863		1134		1626		1772				909		931	
Ę	SEQ	А	S	×	92		93		94		28		95		29		96				30		31	
				Vector	pBluescript		Uni-ZAP XR		Uni-ZAP XR		pCMVSport	3.0	pBluescript	SK-	pSport1		Uni-ZAP XR				pSport1		pSport1	
	ATCC	Deposit	Nr and	Date	209683	03/20/98	209641	02/25/98	97923	03/07/97	209878	05/18/98	209141	16/60/10	209878	05/18/98	106/6	02/26/97	209047	05/15/97	209878	05/18/98	209877	05/18/98
			cDNA	Clone ID	HBMCF37		HFLQB16		HAGFY16		HRACJ35		HAWAZ34		HTWDE26		HMHBN40				HUSIB13		HBAFA02	
			Gene	No.	17		17		17		18		18		19		19				20		21	

		Last	ΑA	d of	ORF	09		248		248		612	_	447		291		184		78		333		164	
		First	AA of	Secreted	Portion	24		61		45		34		56		17		19		22		24		16	
	Last	AA	of	Sig	Pep	23		18		4		33		25		16		18		21		23		15	
	First Last	AA	of	Sig	Pep	ı		1		-		1		-		_		_		-		-		1	
	AA	SEQ		ö	Y	124		125		126		127		128		129		190		130		131		132	
S' NT	jo	First	AA of	Start Signal NO:	Pep	32		126		244		24		28		629		31		19		95		27	
		5° NT	of	Start	Codon	32		126		244		24		28		629		31		61		95		27	
	3, NT	of	Clone	Seq.		1407		1526		1580		2242		2235		2557		1955		1163		1930		881	
	5' NT 3' NT	of	Clone Clone	Seq.		-				41		9		2		260		1		-		28		_	
			Total	K	Seq.	1407		1526		1737		2242		2235		2971		5561		1163		1932		881	
	K	SEQ		ÖN S	X	32		33		34		35		36		37		86		38		39		40	
					Vector	pBluescript	SK-	Uni-ZAP XR		pBluescript		pCMVSport	3.0	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		pSport1		pBluescript	SK-	Uni-ZAP XR	
		ATCC	Deposit	Nr and	Date	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	<i>LL</i> 8607	05/18/98	209877	05/18/98	209580	01/14/98	209877	05/18/98	209877	05/18/98	209877	05/18/98
				cDNA	Clone ID	H2CBT75		HAGDQ42		HBMCJ42		HDPBQ71		HCEIG71		HELHL48		HSKCT36		HISAQ04		HJACB89		HTECC05	
				Gene	No.	22		23		24		25		56		27		27		28		53		30	

		Last	AA	of	ORF	244		55		54		88		19		105		51		119		58		66	
		First	AA of	Secreted	Portion (47		23		56		70		34		21		29		34		94		37	
	Last	AA	o	Sig	Pep	46		22		25		61		33		20		28		33		45		36	
	First Last	AA	o	Sig	Pep	1		-		_				-		-		-		_		-		-	
	AA	SEQ	<u>O</u>	ÖN	Y	133		134		135		136		137		138		139		140		141		142	
5' NT	of	First	AA of	Signal NO:	Pep	217		143		237		148		35		766		131		345		137		66	
		5' NT	of	Start	Codon	217		143		237		148		35		366		131		345		137		66	
	3' NT	oę	Clone	Seq.		1931		1164		1105		1262		517		828		6107		703		689		<i>L</i> 98	
	5' NT 3' NT	o	Clone Clone	Seq.		201						-		_		2		I		I		I		1	
			Total	Z	Seq.	1932		1164		1.105		1262		517		828		2019	i	703		639		298	
	Ł	SEQ	白	ÖN	×	41		42		43		44		45		46		47		48		49		20	
					Vector	Uni-ZAP XR		ZAP Express		Uni-ZAP XR		Uni-ZAP XR		Lambda ZAP	II	ZAP Express		pCMVSport	3.0	pCMVSport	3.0	Uni-ZAP XR		Uni-ZAP XR	
		ATCC	Deposit	Nr and	Date	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98
				cDNA	Clone ID	HBJLF01		HBXGP60		HCE5B20		HCMSQ56		HCNAH57		HCUEP91		HDPCJ91		HDPGK25		HE2DY70		HE2NV57	
				Gene	No.	31		32		33		34		35		36		37		38		39		40	

		Last	AA	ot	R. F.	64		99	Ţ	- 85		55		65		327		68		148		390		86	
		First	AA of /	Secreted	Portion C	22		38		50		70		70	\dashv	24		22		34		27		35	
	Last	AA	Jo	_	Peg	21		37		19		19		19		23		21		33		26		34	
	First Last	AA	jo	Sig	Pep			_		_		_						_		-				-	
	AA	SEQ	0	Ö	≻	143		144		145		146		147		148		149		150		151		152	
S' NT	of	First	AA of	Start Signal NO:	Pep	191		43		4		71		172		232		83		55		73		197	
		5' NT	of	Start	Codon	191		43		4		71		172		232		83		55		73		197	
	3' NT	Jo	Clone	Seq.		1569		1196		945		488		2860		1272		1909		1050		2533		899	
	5' NT 3' NT	of	Clone Clone	Seq.		1		1		-		_		_		93		-						-	
			Total	K	Seq.	1569		1196		945		488		2860		1559		2064		1050		2533		668	
	LN	SEQ	9	ö	X	51		52		53		54		55		99		23		28		29		09	
					Vector	Uni-ZAP XR		Lambda ZAP	II	Lambda ZAP	Ш	Uni-ZAP XR	-	pSport1		pBluescript	SK-	pCMVSport	2.0	pSportl		pCMVSport	3.0	pCMVSport	3.0
		ATCC	Deposit	Nr and	Date	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98	209877	05/18/98
				cDNA	Clone ID	HETBR16		HFXDG13		HFXKY27		HHPEC09		HISAD54		HJBCY35		HKAEA19		HKGDL36		HLDBS43		HLWAD92	
			-	Gene	No.	41		42		43		44		45		46		47		48		49		20	

	First Last	AA AA First Lust	of of AA of AA	Sig Secreted of	Pep Pep Portion ORF	1 22 23 60		1 30 31 392		1 19 20 74		18 91 51 1		1 31 32. 53		1 28 29 102		1 16 17 575		1 16 17 146		1 22 23 643		1 22 23 124	
S' NT	of AA	First SEQ	AA of ID	Signal NO:	Pep Y	92 153		25 154		121 155		138 156		236 157		130 158		191 159		191 191		091 229		392 192	
		5' NT	oę		Codon	92		25		121		138		236		130		161		161		<i>LL</i> 9		392	
	5' NT 3' NT	of	Total Clone Clone	Seq.		1079		1928		781		1194		1677		1237		1934		1958		2729		2444	
	5' NT	of	Clone	Seq.		1		I		I		1		T						-		984		1	
			Total	N	Seq.	1079		1928		781		1194		1677		1237		1934		1958		3300		2444	
	Z	SEQ	Ω	ÖN :	×	19		62		63		64		9		99		<i>L</i> 9		66		89		001	
					Vector	pSport1		Lambda ZAP	II	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		pCMVSport	3.0	pCMVSport	3.0	pCMVSport	3.0	Uni-ZAP XR		Uni-ZAP XR	
		ATCC	Deposit	Nr and	Date	209877	05/18/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98
				cDNA	Clone ID	HLYBIIS		HMEJE05		HNGIX55		HNHEX30		HPJBI33		HRABA80		HRACD80		HRACD80		HSLCX03		HSLCX03	
				Gene	No.	51		52		53		54		55		99		57		57		28		28	

		Last	AA	o	ORF	190		63		117		42		47		45		207		51		20		42	
		First	AA of	Secreted	Portion	56		30		48		21		28		26		31		31		30		30	
	Last	AA	jo	Sig	Pep	25		29		47		70		27		25		30		30		29		29	
	First	AA	of	Sig	Pep	-		_		_ I		I		1				-		-		_		1	
	AA	SEQ		ÖN.	Y	191		162		163		164		165		166		191		193		891		691	
5' NT	oę	First	AA of	Signal NO:	Pep	122		213		173		28		55		125		157		157		94		102	
		5' NT	of	Start	Codon	122		213		173		28		55		125		151		157		94		102	
	3' NT	of	Clone	Seq.		1797		1373		1579		1028		3674		2797		2703		2709		742		1825	
	S' NT 3' NT	oę	Clone Clone	Seq.		92		_		1		1		I		I		I		_				-	
			Total	Ľ	Seq.	1797		1373		1579		1028		3674		2797		2703		2709		742		1825	
	Ľ	SEQ	白	SO.	×	69		70		71		72		73		74		75		101		9/		11	
					Vector	Uni-ZAP XR		Uni-ZAP XR	- 1-	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		pBluescript	SK-	pBluescript	SK-	pBluescript	SK-	Uni-ZAP XR		Uni-ZAP XR	
		ATCC	Deposit	Nr and	Date	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98	209889	05/22/98
				cDNA	Clone ID	HT5GJ57	:	HTACS42		HTEKE40		HTOBX69		HUVE077		H2CBG48		H2CBU83		H2CBU83		HAPNY94		HBJHZ58	
				Gene	No.	59		09		61		62		63		64		65		65		99		67	

NT SEQ ID	NT SEQ ID Total	NT SEQ ID Total	Total		5' NT 3' NT of Olone Clone	5' NT 3' NT of of Clone Clone	S' NT of	of of 5' NT SEQ AA AA Single Clone of AA of ID of of Single Clone of AA of ID of of of Single Clone of Single Clone of Single Si	of AA First Last irst SEQ AA AA AA Aof ID of of of	First L AA of	Last AA of	First AA of	Last AA
Nr and Date Vector	Vector		į ×	Seq.	Seq.	seq.	Codon	Start Signal INC. Sign Codon Pep Y Pep	- K		Nep Pep	Y Pep Pep Portion	ORF
209889 Uni-ZAP XR 05/22/98	Uni-ZAP XR	_	78	1674	1	1668	67	67	170	-	18	19	55
209889 pCMVSport 05/22/98 3.0			79	79 2191	291	2191	460	460	171	1	24	25	108
209889 pSport1 05/22/98	pSport		08	1335	1	1335	43	43	172	_	15	16	20
209889 pCMVSport 05/22/98 3.0			81	1867	415	1867	103	103	173	_	21	22	996
209889 pCMVSport 05/22/98 3.0			102	1722	1	1722	59	59	194	-	21	22	319

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

10

15

20

25

30

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further

below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

5

15

20

25

30

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits.

Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed

sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

5

10

15

20

25

30

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the

191

information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polynucleotide and Polypeptide Variants

5

10

15

20

25

30

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown inTable 1, the ORF (open reading frame), or any fragement specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the presence invention can be determined conventionally using known computer programs. A preferred method for determined the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignement of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95%

"identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N-and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the

195

query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

5

10

15

20

25

30

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the Nterminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and Ctermini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter

196

the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

10

15

20

25

30

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See,

197

Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

5

10

15

20

25

30

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used.

(Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

5

10

15

20

25

30

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

15

20

25

30

5

10

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-

1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

10

15

20

25

30

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and

201

alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

5

10

15

20

25

30

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

202

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

20

25

30

5

10

15

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous

203

functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

5

10

15

20

25

30

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See,

D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

15

20

25

30

10

5

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The

expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

5

10

15

20

25

30

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography,

phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence). and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al.,

207

Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

5

10

15

20

25

30

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however,

polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques." Pergamon Press, New York (1988).

5

10

15

20

25

30

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression,

209

chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

5

10

15

20

25

30

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

5

10

15

20

25

30

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA

211

antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

5

10

15

20

25

30

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (1251, 1211), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety

needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

5

10

15

20

25

30

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also

be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

5

10

15

20

25

30

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia). ataxia telangiectasia, common variable immunodeficiency,

PCT/US99/15849

5

10

15

20

25

30

Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

25 **Hyperproliferative Disorders**

5

10

15

20

30

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

5

10

15

20

25

30

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the

217

present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae. Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, 5 Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye 10 infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide 15 or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive 20 bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis. Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, 25 Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, 30 endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis,

opportunistic infections (c.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

5

10

15

20

25

30

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See,

219

Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

5

10

15

20

25

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

5

10

15

20

25

30

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable

of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

5

10

15

20

25

30

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a

candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

5

10

15

20

25

30

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95%

identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

5

10

15

20

25

30

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

5

10

15

20

25

30

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

5

10

15

20

25

30

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least

226

one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

5

10

15

20

25

30

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

5

10

15

20

25

30

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in

the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

5

10

15

20

25

30

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide

comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

5

10

15

20

25

30

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino

acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

5

10

15

20

25

30

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is

231

expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

20

25

30

5

10

15

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	Vector Used to Construct Library	Corresponding Deposited
	<u>Plasmid</u>	
	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
5	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSport 2.0	pCMVSport 2.0
	pCMVSport 3.0	pCMVSport 3.0
10	pCR [©] 2.1	pCR®2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain

XL-1 Blue. Vector pCR[®]2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

5

10

15

20

25

30

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ³²P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for

bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

5

10

15

20

25

30

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with

phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

5

10

15

20

25

30

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHybTM hybridization solution (Clontech) according to

236

manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

5 Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions: 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

10

15

20

25

30

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc.,

Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain

M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

5

10

1.5

20

25

30

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., supra). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., supra).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250

mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains:

1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

5

10

15

20

25

30

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM

239

Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

5

10

15

20

25

30

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium

240

acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A₂₈₀ monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

10

15

20

25

30

5

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in

241

Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

5

10

15

20

25

30

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. E. coli HB101 or other suitable E. coli hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five μg of a plasmid containing the polynucleotide is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μg of BaculoGold™ virus DNA and 5 μg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded

in a 35 mm tissue culture plate with 1 ml Grace's medium without scrum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

5

10

15

20

25

30

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (c.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μCi of ³⁵S-methionine and 5 μCi ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

5

10

15

20

25

30

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991): Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a

5

10

15

20

25

in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μCi of ³⁵S-methionine and 5 μCi ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

5

10

15

20

25

in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (c.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μ Ci of ³⁵S-methionine and 5 μ Ci ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

5

10

15

20

25

30

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991): Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a

chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

5

10

15

20

25

30

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide.

Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. E. coli HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., supra). The plasmid pSV2-neo

that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μM, 2 μM, 5 μM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

15 Example 9: Protein Fusions

5

10

20

25

30

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These

primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

15 Human IgG Fc region:

5

10

20

25

30

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGC
CCAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAAA
CCCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGT
GGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG
ACGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA
CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT
GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA
ACCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG
GTCAGCCTGACCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG
GTCAGCCTGACCTGCCCCCATCCCGGGAGAACAACTACAAGACCACGCCT
CCCGTGCTGGACTCCGACGCCTCCTTCTTCCTCTACAGCAAGCTCACCGTG
GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCA
TGAGGCTCTGCACAACCACTACACGCAGAAGACCTCCCCTGTCTCCCGG
GTAAATGAGTGCGACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

5

10

15

20

25

30

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 μg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is

248

possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

25

30

20

5

10

15

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

5

10

15

20

25

30

First. dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2 x 10⁵ cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl2 (anhyd); 0.00130 $mg/L CuSO_4-5H_2O$; 0.050 mg/L of $Fe(NO_3)_3-9H_2O$; 0.417 mg/L of $FeSO_4-7H_2O$; 311.80 mg/L of Kcl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄-H₂0; 71.02 mg/L of Na_2HPO4 ; .4320 mg/L of ZnSO₄-7H₂O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-10 Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂0; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H₂0; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂0; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-15 Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalainine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tryrosine-2Na-2H20; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of 20 Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of 25 Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for 30 endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

15 Example 12: Construction of GAS Reporter Construct

5

10

20

25

30

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferonsensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2,

252

Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

5

10

15

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

253

	Ligand	tyk2	JAKs Jakl	Jak2	Jak3	<u>STATS</u>	GAS(elements) or ISRE
5	<u>IFN family</u> IFN-a/B	+	+	-	-	1,2,3	ISRE
	IFN-g II-10	+	+ ?	+ ?	-	1 1,3	GAS (IRF1>Lys6>IFP)
	•	·	•	·		-,-	
10	gp130 family IL-6 (Pleiotrophic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
10	Il-11(Pleiotrophic)	?	+	?	?	1,3	•
	OnM(Pleiotrophic)	?	+	+	?	1,3	
	LIF(Pleiotrophic)	?	+	+	?	1,3	
	CNTF(Pleiotrophic)	-/+	+	+	?	1,3	
15	G-CSF(Pleiotrophic)	?	+	?	?	1,3	
	IL-12(Pleiotrophic)	+	-	+	+	1,3	
	g-C family						
	IL-2 (lymphocytes)	_	+	_	+	1,3,5	GAS
20	IL-4 (lymph/myeloid)	٠.	+	_	+	6	$GAS_{\cdot}(IRF1 = IFP >> Ly6)(IgH)$
20	IL-7 (lymphocytes)	_	+	_	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	5	GAS
	IL-13 (lymphocyte)	_	+	?	?	6	GAS
	IL-15	?	+	?	+	5	GAS
25		-					
	gp140 family						
	IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	5	GAS
	GM-CSF (myeloid)	-	-	+	-	5	GAS
30							
	Growth hormone fam	<u>ily</u>					·
	GH	?	-	+	-	5	•
	PRL	?	+/-	+	-	1,3,5	
	EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
35							
	Receptor Tyrosine Ki						CAC (IDE1)
	EGF	?	+	+	-	1,3	GAS (IRF1)
	PDGF	?	+	+	-	1,3	CAS (not IDE1)
40	CSF-1	?	+	+	-	1,3	GAS (not IRF1)
40							

5

10

15

30

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCCGAAATGATTTCCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATITCCCCGAAATCTAGATITCCCCGAAATGATITCCCCGAAA
TGATITCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCG
CCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCT
CCGCCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGC
TCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT
25 AGGCTTTTGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol

acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and Xhol, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Nco vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, Il-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

25

30

5

10

15

20

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells

(ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml genticin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

5

10

15

20

25

30

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10⁷ per transfection), and resuspend in OPTI-MEM to a final concentration of 10⁷ cells/ml. Then add 1ml of 1 x 10⁷ cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

5

10

15

20

25

30

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

The above protocol may be used in the generation of both transient, as well as, stable transfected cells, which would be apparent to those of skill in the art.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell

258

Growth & Differentiation, 5:259-265) is used. First, harvest 2x10e⁷ U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM Na₂HPO₄.7H₂O, 1 mM MgCl₂, and 675 uM CaCl₂. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

5

10

15

20

25

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting 1×10^8 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in the 96-well plate (or 1×10^5 cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (carly growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

5

10

15

20

25

30

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

- 5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)
- 5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heatinactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine

growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

5

10

15

20

25

30

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as $5x10^5$ cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF-κB (Nuclear Factor κB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-κB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- κB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κB is retained in the cytoplasm with I-κB (Inhibitor κB). However, upon stimulation, I- κB is phosphorylated and degraded, causing NF- κB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κB include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

5

10

20

25

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-kB promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF-kB would be useful in treating diseases. For example, inhibitors of NF-kB could be used to treat those diseases related to the acute or chronic activation of NF-kB, such as rheumatoid arthritis.

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

15 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)
Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCC
ATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGA
CTAATTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTA
TTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAA
GCTT:3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-30 promoter plasmid (Clontech) with this NF-kB/SV40 fragment using XhoI and

HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-κB/SV40/SEAP cassette is removed from the above NF-κB/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF-κB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF-kB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1,1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

5

10

15

20

25

30

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense $15~\mu l$ of 2.5x dilution buffer into Optiplates containing $35~\mu l$ of a supernatant. Seal the plates with a plastic sealer and incubate at $65^{\circ}C$ for 30~min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 µl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 µl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

5

Reaction	Buffer	Formu	lation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80 .	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants

which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

5

10

15

20

25

30

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.; catalog no. F-14202), used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To load the cells with fluo-4, 50 ul of 12 ug/ml fluo-4 is added to each well. The plate is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-4. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm;

and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca⁺⁺ concentration.

5 Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

10

15

20

25

30

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St.

Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum. rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

5

10

15

20

25

30

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from Boeheringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16.000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

267

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂₊ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

5

10

15

20

30

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavadin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phospotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

25 Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine

5

10

15

20

25

phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1

and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place

of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (lug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

5

10

15

20

25

30

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro. VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv.

et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate I hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard

5

10

15

20

25

30

271

curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

5

10

15

20

25

30

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semipermeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

5

10

15

20

25

30

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's

solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

5

10

15

20

25

30

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical

274

compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

20

25

30

5

10

15

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

5

10

15

20

25

30

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector.

The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

5

10

15

20

25

30

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 27: Method of Treatment Using Gene Therapy - In Vivo

Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479, Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord. 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).

277

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

5

10

15

20

25

30

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection

into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

5

10

15

20

25

30

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle in vivo is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

279

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

Example 28: Transgenic Animals.

5

10

15

20

25

30

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (i.e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., Science 259:1745 (1993); introducing nucleic acid constructs into embryonic pleuripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-

mediated gene transfer (Lavitrano et al., Cell 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

5

10

15

20

25

30

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al., Science 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA

expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

25 Example 29: Knock-Out Animals.

5

10

15

20

30

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (E.g., see Smithies et al., Nature 317:230-234 (1985); Thomas & Capecchi, Cell 51:503-512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding

282

regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e.g., see Thomas & Capecchi 1987 and Thompson 1989, supra). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

5

10

15

20

25

30

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e.g., knockouts) are administered to a patient in vivo. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

5

10

15

20

25

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

A. The indications made below relate to the microorganism referred to in the description on page 178		
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Collect	etion .	
Address of depositary institution (including postal code and country 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	y)	
Date of deposit	Accession Number	
May 18, 1998	209878	
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet	
D. DESIGNATED STATES FOR WHICH INDICATION	S ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave bl	lank if not applicable)	
The indications listed below will be submitted to the Internation Number of Deposit")	al Bureau later (specify the general nature of the indications e.g., "Accession .	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer Authorized officer Literature of Division 708-808-2082	Authorized officer .	

A. The indications made below relate to the microorganis on page, line	ism referred to in the description N/A .
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Cultur	re Collection
Address of depositary institution (including postal code of 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	and country)
Date of deposit	Accession Number
February 26, 1997	97898
C. ADDITIONAL INDICATIONS (leave blank if not	applicable) This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDIC	CATIONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATION	NS (leave blank if not applicable)
	nternational Bureau later (specify the general nature of the indications e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international applica	11
Authorized officer Externational Division 705-005-2022	Authorized officer

A. The indications made below relate to the microorganism referred to in the description		
on page 179 line	N/A .	
B. IDENTIFICATIONOF DEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Colle	ction	
Address of depositary institution (including postal code and count	(ry)	
10801 University Boulevard		
Manassas, Virginia 20110-2209 United States of America		
·	·	
Date of deposit	Accession Number	
May 15, 1997	209044	
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet	
	·	
D. DESIGNATED STATES FOR WHICH INDICATION	NS ARE MADE (if the indications are not for all designated States)	
·		
·	•	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)		
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")		
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer	Authorizedofficer	
laternational multion		
Transport of the second		

accoción iciaioto inc.	incroor Barnsin receiv	ed to in the description
179	,line	N/A
ONOFDEPOSIT		Further deposits are identified on an additional sheet
titution American Ty	ype Culture Colle	ction
		·
loulevard a 20110-2209	postal code and count	(קי
		Accession Number
August 28, 1997		209225
INDICATIONS (leav	e blank if not applicabl	This information is continued on an additional sheet
RNISHING OF IND	DICATIONS (leave b	NS ARE MADE (if the indications are not for all designated States) Plank if not applicable) The indications e.g., "Accession
ceived with the internat	ional application	For International Bureau use only This sheet was received by the International Bureau on: Authorized officer
	DNOFDEPOSIT Initiation American Tyle institution (including Boulevard a 20110-2209 merica August 28, 1997 INDICATIONS (leave of the content of the conten	DNOFDEPOSIT Initiation American Type Culture College In institution (including postal code and count Boulevard A 20110-2209 Merica August 28, 1997 INDICATIONS (leave blank if not applicable) STATES FOR WHICH INDICATION RNISHING OF INDICATIONS (leave blank if not applicable) below will be submitted to the Internation ecceiving Office use only ceived with the international application

A. The indications made below relate to the microorganism referre		
on page 180, line	N/A	
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Collect	ction .	
Address of depositary institution (including postal code and country	7)	
10801 University Boulevard		
Manassas, Virginia 20110-2209 United States of America		
United States of America		
Date of deposit	Accession Number	
March 7, 1997	97923	
C. ADDITIONAL INDICATIONS (leave blank if not applicable	e) This information is continued on an additional sheet	
•		
	\cdot	
D. DESIGNATED STATES FOR WHICH INDICATION	NS ARE MADE (if the indicutions are not for all designated States)	
·		
	,	
E. SEPARATE FURNISHING OF INDICATIONS (leave b		
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")		
	,	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer Retailer	Authorized officer	
Authorized officer Maintain Phylolega	7.44.10.12.00	
жкатепроод Пур <mark>ојор</mark> 78 %-оподере	· .	

A. The indications made below relate to the microorganism reference on page180, line	ed to in the description N/A .
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Collect	ction .
Address of depositary institution (including postal code and country 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	(ب ر
Date of deposit	Accession Number
May 22, 1997	209071
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION	NS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS (leave b	lank if not applicable)
The indications listed below will be submitted to the Internation Number of Deposit")	
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer (avarrantemn) Division 703-003-0032 micry and interpretations	Authorized officer

PCT/US99/15849

A. The indications made below relate to the microorganism referred to in the description on page 180 Inc		
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Collection		
Address of depositary institution (including postal code and country 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	y)	
Date of deposit	Accession Number 209626	
February 12, 1998		
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet	
·		
D. DESIGNATED STATES FOR WHICH INDICATION	NS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave to	olank if not applicable)	
The indications listed below will be submitted to the Internation Number of Deposit")	nal Burcau later (specify the general nature of the indications e.g., "Accession	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer Matth 1988 to 1988	Authorized officer	

A. The indications made below relate to the microorganism referred to in the description on page 181 N/A		
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Collect	etion	
Address of depositary institution (including postal code and country 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	y)	
Date of deposit	Accession Number	
March 20, 1998	209683	
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet	
D. DESIGNATED STATES FOR WHICH INDICATION		
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)		
The indications listed below will be submitted to the Internation Number of Deposit")	nal Bureau later (specify the general nature of the indications e.g., "Accession	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer CKCHC Comp. Tolkielon 705-005-007	Authorized officer	

A. The indications made below relate to the microorganism referred to in the description on page		
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Colle	ction	
Address of depositary institution (including postal code and count 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	n))	
Date of deposit	Accession Number .	
February 25, 1998	209641	
C. ADDITIONAL INDICATIONS (leave blank if not applicable	(e) This information is continued on an additional sheet	
D. DESIGNATED STATES FOR WHICH INDICATIO	NS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave	blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")		
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer	Authorized officer .	

A. The indications made below relate to the microorganism referred to in the description on page 181 line N/A		
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Collection		
Address of depositary institution (including postal code and country 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	v)	
Date of deposit	Accession Number	
July 9, 1997	209141	
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet	
D. DESIGNATED STATES FOR WHICH INDICATION	S ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)		
The indications listed below will be submitted to the Internation Number of Deposit")	al Bureau luter (specify the general nature of the indications e.g., "Accession	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer 70% (1996)	Authorized officer	

A. The indications made below relate to the microorganism referred to in the description on page		
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Collection		
Address of depositary institution (including postal code and country 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	,	
Date of deposit	Accession Number	
February 26, 1997	97901	
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet	
D. DESIGNATED STATES FOR WHICH INDICATION	IS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)		
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")		
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application	This sheet was received by the International Bureau on:	
Authorized officer	Authorized officer	

A. The indications made below relate to the microorganism refers on page 181, line	ed to in the description N/A .
B. IDENTIFICATIONOF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and count 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	(v)
Date of deposit	Accession Number
May 15, 1997	209047
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION	·
E. SEPARATE FURNISHING OF INDICATIONS (leave	
The indications listed below will be submitted to the Internation Number of Deposit")	nal Bureau later (specify the general nature of the indications e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer .

A. The indications made below relate to the microorganism referred on page181, line	ed to in the description N/A
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Collect	ction
Address of depositary institution (including postal code and country 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	y)
Date of deposit	Accession Number
May 18, 1998	209877
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION	NS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS (leave b	olank if not applicable)
The indications listed below will be submitted to the Internation Number of Deposit")	nal Bureau later (specify the general nature of the indications e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer Science Weiters Weiters Division 783-505-6832	Authorized officer .

A. The indications made below relate to the microorganism referr on page182, line	red to in the description N/A .
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture College	ction
Address of depositary institution (including postal code and count 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	η)
Date of deposit	Accession Number
January 14, 1998	209580
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION E. SEPARATE FURNISHING OF INDICATIONS (leave)	
	nal Bureau later (specify the general nature of the indications e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

A. The indications made below relate to the microorganism referred on page 185, line	ed to in the description N/A .
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Collect	etion
Address of depositary institution (including postal code and countre 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	y)
Date of deposit	Accession Number 209889
May 22, 1998	
C. ADDITIONAL INDICATIONS (leave blank if not applicable	This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION	·
E. SEPARATE FURNISHING OF INDICATIONS (leave b	
The indications listed below will be submitted to the Internation Number of Deposit")	al Bureau later (specify ine general nature of the thalcallors e.g., Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

WO 00/04140 PCT/US99/15849

299

What Is Claimed Is:

5

10

15

- 1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
 - (f) a polynucleotide which is a variant of SEQ ID NO:X;
 - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
- (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
- 25 (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

- 2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
- The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 15 5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
 - 7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.
 - 8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.
 - 9. A recombinant host cell produced by the method of claim 8.
 - 10. The recombinant host cell of claim 9 comprising vector sequences.

15

20

- 11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
- (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (g) a variant of SEQ ID NO:Y;
 - (h) an allelic variant of SEQ ID NO:Y; or
 - (i) a species homologue of the SEQ ID NO:Y.
 - 12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.
 - 13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.
- 25 14. A recombinant host cell that expresses the isolated polypeptide of claim 11.
 - 15. A method of making an isolated polypeptide comprising:
 - (a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and
 - (b) recovering said polypeptide.

- 16. The polypeptide produced by claim 15.
- 17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.
- 18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and
 - (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.
- 19. A method of diagnosing a pathological condition or a susceptibility to
 15 a pathological condition in a subject comprising:
 - (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
 - (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.
 - 20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:
 - (a) contacting the polypeptide of claim 11 with a binding partner; and
- (b) determining whether the binding partner effects an activity of thepolypeptide.
 - 21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.
- 22. A method of identifying an activity in a biological assay, wherein the method comprises:
 - (a) expressing SEQ ID NO:X in a cell;
 - (b) isolating the supernatant:

- (c) detecting an activity in a biological assay; and
- (d) identifying the protein in the supernatant having the activity.
- 23. The product produced by the method of claim 20.

1. .

1

Sequence Listing

```
<110> Human Genome Sciences, Inc., et al.
 <120> 71 Human Secreted Proteins
<130> PZ030PCT
 <140> Unassigned
 <141> 1999-07-14
 <150> 60/092,956
 <151> 1998-07-15
 <150> 60/092,921
 <151> 1998-07-15
 <150> 60/092,922
 <151> 1998-07-15
 <160> 262
 <170> PatentIn Ver. 2.0
 <210> 1
 <211> 733
 <212> DNA
 <213> Homo sapiens
 <400> 1
 gggatccgga gcccaaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg
                                                                           60
                                                                          120
 aattcgaggg tgcaccgtca gtcttcctct tcccccaaa acccaaggac accctcatga
. tctcccggac tcctgaggtc acatgcgtgg tggtggacgt aagccacgaa gaccctgagg
                                                                          180
 tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg
                                                                          240
 aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact
                                                                          300
 ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca acccccatcg
                                                                          360
 agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc
                                                                          420
 catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct
                                                                          480
 atccaagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga
                                                                          540
 ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag ctcaccgtgg
                                                                          600
                                                                          660
 acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc
 acaaccacta cacgcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc
                                                                          720
                                                                          733
 gactctagag gat
 <210> 2
 <211> 5
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> Site
  <222> (3)
 <223> Xaa equals any of the twenty naturally ocurring L-amino acids
  <400> 2
  Trp Ser Xaa Trp Ser
```

WO 00/04140 PCT/US99/15849

<210> 3 <211> 86 <212> DNA <213> Homo	sapiens					
	atttccccga ctgccatctc		tccccgaaat	gatttccccg	aaatgatttc	60 86
<210> 4 <211> 27 <212> DNA <213> Homo	o sapiens					
<400> 4 gcggcaagct	ttttgcaaag	cctaggc				27
<210> 5 <211> 271 <212> DNA <213> Homo	sapiens					
aaatatctgo gcccctaact ttatgcagag	c ccccgaaatc c catctcaatt c ccgcccagtt g gccgaggccg c ctaggctttt	agtcagcaac ccgcccattc cctcggcctc	catagtcccg tccgccccat tgagctattc	ggctgactaa	tttttttat	60 120 180 240 271
<210> 6 <211> 32 <212> DNA <213> Home						
<400> 6 gcgctcgag	g gatgacagcg	atagaacccc	gg			32
<210> 7 <211> 31 <212> DNA <213> Hom	o sapiens					
<400> 7 gcgaagctt	c gegaeteece	ggateegeet	C			. 31
<210> 8 <211> 12 <212> DNA <213> Hom	o sapiens					
<400> 8 ggggacttt						12

```
<210> 9
<211> 73
<212> DNA
<213> Homo sapiens
<400> 9
gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg
                                                                         60
                                                                         73
ccatctcaat tag
<210> 10
<211> 256
<212> DNA
<213> Homo sapiens
<400> 10
                                                                         60
ctcgagggga ctttccggg gactttccgg ggactttcca tctgccatct
caattagtca gcaaccatag tcccgccct aactccgccc atcccgcccc taactccgcc
                                                                        120
cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg cagaggccga
                                                                        180
ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg
                                                                        240
                                                                        256
cttttgcaaa aagctt
<210> 11
<211> 1113
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (393)
<223> n equals a,t,g, or c
<400> 11
gatgctcctt tagcttggag gagtttgtta ttacccacct tctgaagcct acttctgtca
                                                                         60
attcatccaa ctcattctca gtccagtttt gtttccttgc tggtgaggag ttgtgatcct
                                                                        120
ttggaggaga agaggcattc tggtttttgg aatttttagc cattttgctc tggtttcttc
                                                                        180
ccatctttgt ggatttatct acctttcatc ttcaatgtta gtgacctatg gatggggtct
                                                                        240
ctgagtggat gtgctcttcc tttctgtttg twagttttct ttctaacagt tagcccctct
                                                                        300
gctgtaggtc tgctggaktt tgctggaggt ccactccaga ccctgtttgc ctgggtatca
                                                                        360
ccagtggagg ctgcagaaca gcaaagattg ctncctgttc tttcctctgg aagcttcgtc
                                                                        420
 tcagagggca cctgccagat gccagccaga gctctcctgt atgaggtgtc tgttggccca
                                                                        480
 tactgggaga ttcctcccag tcaggataca aggaggtcag ggacctactt gaggaggcag
                                                                        540
 totgaccott agcagaggtt gaacactgtg ctaggaggtc ctctgctctt ttcagagctg
                                                                        600
 tcaggcgggg cgtataagtc tgctgaatct gtgtccgcag ccaccccttc ccccaggtgc
                                                                        660
                                                                        720
 tctgtcccag ggagatgggg gttttatttt taagtcccca actggggctg ctgccttttt
                                                                        780
 ttcagagatg ccctgcccag agaggagaaa tctggcagtc tggcctcaga ggccttgctg
 agctgccgtt ggctccaccc agttcaaact tcccaagggg ctttgtttat actgtaaggg
                                                                        840
 gaaaaccgcc tactcgagcc tcatcaatgg cagacacccc tccccgcgcc aagcttaagt
                                                                        900
gtcccaggtt gatctcagac tgctgctgtg ctggcagtga gaatttcaag ccagtggatc
                                                                        960
 ttagtttgct gtgctctgtg ggggtgggac ccattgaacc agactactcg gctccctggc
                                                                       1020
 ttcagccccc tttccaggag agtgaagggt tctgtctcat tggcattcca ggagtcactg
                                                                       1080
                                                                       1113
 gtgtatggaa aaaaaaaaa aaagggcggc cgc
```

Δ

```
<211> 983
<212> DNA
<213> Homo sapiens
<400> 12
ggcacgaggg cagctgcaga gctccaggtt tctctgccca caagggcagg ggctgcccct
                                                                         60
                                                                        120
cgcccaggat gactctgcct tccagagcct tggcctccct gggggtggga gtgtgggga
                                                                        180
tgctaaggtt aaatcaggtc acagtaagtt gtgggggcag caggtggagc agcagagtgg
                                                                        240
cactgggage tttctcttgg gtgtgcggtg tggccttggt tctgcagcca tcaggtgggg
                                                                        300
gcttgggact gacttctcct tctgaaggat gctgggaagg tgagctggct ttggcagtgc
ttagagetee ggggggttee ceeteetaga acatgeaage teteacaceg gtgegteate
                                                                        360
                                                                        420
atcacaccca tcatcaagcc cacagtggta tactgaacac ctgccccaca aagacggtgg
                                                                        480
actgctctca gaggagcccc atgaaccacc gatggttaca actatccaat gcctgatggc
agacagecag gecaaceteg gettecaete tetetteete accetacaat cagecaaagt
                                                                        540
gacctgagtc atgtagtgtg aagttgcttt ctgctttctc ttgtttgtgc ttttgctgtt
                                                                        600 -
tettetgece catactttgt taactecatg agttaaatge tacceatttt cecagacaag
                                                                        660
tgctgcttct gcaaggaaac ccttcctgat ccccaccta tctgaaaagt acctctccag
                                                                        720
cttgcttctt cagggtgctg agcgttcctt cccagcctgt catcaccttc ctccatacgc
                                                                        780
tatggtgtgt tcctgtcttc tctagtcttg tcctcttttt tctgttagat tgtagctcct
                                                                        840
tgctgacagg aaccacgcct gctccagctt catacctccc actgctacag cacagaacct
                                                                        900
gcttctcaga cttacagcaa atgtttgttt gctgaatgaa ttaattaaag ataaagcaaa
                                                                        960
                                                                        983
aaaaaaaaa aaaaaaactc gag
<210> 13
<211> 973
<212> DNA
<213> Homo sapiens
<400> 13
                                                                         60
ggcacgagcc cageggaagc caagecacca ggeeeeeeag egteeaegeg gageatgaae
                                                                        120
attgaggatg gcgcgtgccc gcggctcccc gtgccccccg ctgccgcccg gtaggatgtc
                                                                        180
ctggccccac ggggcattgc tettectetg getettetee ceacceetgg gggccggtgg
                                                                        240
aggtggagtg gccgtgacgt ctgccgccgg agggggctcc ccgccggcca cctcctgccc
                                                                        300
cgtggcctgc tcctgcagca accaggccag ccgggtgatc tgcacacgga gagacctggc
cgaggtccca gccagcatcc cggtcaacac gcggtacctg aacctgcaag agaacggcat
                                                                        360
ccaggtgatc cggacggaca cgttcaagca cctgcggcac ctggagattc tgcagctgag
                                                                        420
caagaacctg gtgcgcaaga tcgaggtggg cgccttcaac gggctgccca gcctcaacac
                                                                        480
                                                                        540
gctggagctt tttgacaacc ggctgaccac ggtgcccacg caggccttcg agtacctgtc
                                                                        600
caagctgcgg gagctctggc tgcggaacaa ccccatcgag agcatcccct cctacgcctt
                                                                        660
caaccgcgtg ccctcgctgc ggcgcctgga cctgggcgag ctcaagcggc tggaatacat
                                                                        720
ctcggaggcg gccttcgagg ggctggtcaa cctgcgctac ctcaacctgg gcatgtgcaa
                                                                        780
cctcaaggac atccccaacc tgacggccct ggtgcgcctg gaggagctgg agctgtcggg
caaccggctg gacctgatec gecegggete ettecagggt etcaccagee tgegcaaget
                                                                        840
                                                                        900
gtggctcatg cacgcccagg tagccaccat cgagcgcaac gccttcgacg acctcaagtc
                                                                        960
getggaggag etcaacetgt eccacaacaa ectgatgteg etgeeccaeg acetetteae
                                                                        973
gcccctgcac cgc
<210> 14
 <211> 1458
 <212> DNA
 <213> Homo sapiens
 <400> 14
ccacgcgtcc gggaattttc aaaagatcca aacagagact tcctgcatct tctgcctttc
                                                                          60
                                                                         120
 caacagaagc ggtgatcgtc taagtatgag cctgtggctt cctttgtgca tttgagcatg
ctgtaattaa gatgagatca gtttcttaga aaaagctttc ctgaatccct ctgacgttgc
                                                                         180
```

				acadadcatc	tataatccaa	240
ctgggatctt	tctgttgatt	cgtcttttct	ggagaccggg	acagageace	agacattaca	300
ggaagttagt	cctctggcct	caattctgtt	grggargrgc	aytyataayo	tatattaa	360
+t-00000	gatgeetagt	teatagette	ctggctgttt	Cyccecee	og co o og cang	420
	ccaactcaaa	asataaata	EEggeggege	ccccgcggcc	55	480
	cctccadatt	chagacatga	qaqqccgcgcg	CCCCACCACC	99-33	
	taatamoato	tttaattaga	tcacagcall	gaatttaaaa	0000000	540
	ataaaacata	agacaccodd	aatgagygay	gaggaagaca	900909	600
adayadaycc	attanactat	agccgatgca	tttgtcagga	aatacagcac	tttgtcttaa	660 [.]
CtCtttaaac	cityagetee	ggcgcagtgg	ctcacaccta	taatcccagc	actttgggag	720
gaaaacaagg	aaggaggccg	tgaggtgggg	agtatgagac	caccctgact	aacatggaga	780
gccgaggcgg	gcggatcacc	tgaggtgggg	ageaegage	grtgcgcatg	cccataatcc	840
gaccctgtct	ctactaaaag	tacagaatta	gccgggcgcg	ctcaccaca	gaggttgcag	900
cagctactga	ggagacttga	ggtaggagaa	tcacttgaac	cccagcggcg	taggteteaa	960
·	tagagagagt	acactccage.	Ctdddcaaya	agagegaaa	0999000	1020
	~~~~~~~~~	Jasagagtaat	ttacaacyaa	ggaaaaaaa		1080
	atatasacac	tetectgatg	tcacagtggc	Lycycycoc	-99999-999-	1140
	~~~~~~~~~	cctaacccta	CCCCCCCCCCC	CCCGCCCCC		
	aggrataged	addcctccdc	tegtillaya	Cacgaacce		1200
	catatacact	aratataa	agcagaatyc	Cigcccccc	09900	1260
cacaacacg	thetacacage	tattattat	tetttaccaa	ttttgggaat	aaaaggtgtt ttcaacttgt	1320
ttgtaccata	CEGLAdalla	caccacccac	acttctattt	aaatttccag	ttcaacttgt	1380
ttggttattt	aatataataa	gagetgetaa	actitigeto	taaaaaaaaa	aaaaaaaaa	1440
aaatgttttt	attgtgcata	aatacatact	aacyctgacc		aaaaaaaaa	1458
aaaaaaaaaa	aaaaaaa					

<210> 15 <211> 2005 <212> DNA

<213> Homo sapiens

		•				
<400> 15				acaaacacat	ccccgcgacg	60
ggttgctggc	ccaggtgagc	gggcgcgctg	gcccaggcga	tectaagge	cagccgggaa	120
gcgctgcctg	cccgaggcgg	ttcacgtaaa	gacagcgaga	cccacaccac	tacctagege	180
ggaggcgtgg	atatggagct	ggctgctgcc	aagtccgggg	cadadaccca	tagagcccgg	240
gtcctgggga	ctctgtgggg	acgcgccccg	egeegeggee	ctcctactcc	tecteatoge	300
cgctgcgcgc	atggccctgc	tctcgcgccc	egegeteace	tagadageca	ccctgaagac	360
cgctgttgtc	aggtgccagg	agcaggccca	gaccaccyac	accaccttag	acctcctggg	420
catccggaac	ggcgttcata	agatagacac	gtacctgaac	tetaageett	tcccacatta	480
		antataaato	cagtgacgga	CCCaagecee	0000	540
	anatacacac	caaataaata	EddCfCfcca	Ctgtttggtg	,	600
	aattacataa	caaagtgttg	caaccaacac	gacaggegee	~~	660
L	nagaatgact	ataataaaa	attccagtat	cyceceeca	4 545555	720
	222222424	ractaactca	dcatdtttay	gcacgcgaaa	0440-5-55	780
	ascsatatta	tacatttagg	ttgtaaacca	Latteggata	900	840
	tatasttata	aagaaaaac	tgatctttaa	aggagacgcc	940490	900
	astaassass	cataaccttt	gacaaataac	Caatgeeee		960
	++++a+a222	ggattattt	gagaccttaa	aacuucccas		1020
L	224422222	antnannnan	acadcdayyy	gagggcacgc	••9••	1080
		ctcccttact	tagtatqcca	aacyccccga	00000	1140
	++~+++=~~~	- dadaattttu	aaaaqaqqaa	Lacacaacca		1200
		adatcaaata	taaaatttat	Cacaacgccc	5000	1260
		aaattatcac	ttacaaqtat	Legiculation	• 5	1320
		- addaacddac		Lacteude		1380
	~+~~~~~	atatottott	CCCCCCCAGC	Cegeegagea	0005-5	1440
		- cttmcaagga	caaddiiii	. acceces	3-	1500
		tratagggag	caaccccagg	acagecaage	9	1560
		- atacaattca	daadcatage	, caccecati		1620
		· Efftactata	acctttqcc	, agttugugue		1680
taatgagata	aatgaaaaga	tatttamcct	: aatatatcaa	ggcactattt	gctgttatgc	1080

6	
totratagar trittaaaga cigtaaccit	1740
tttgttattt atttcccagc acttgttcct tattgtagat tttttaaaga ctgtaacctt	1800
	1860
	1920
	1980
ttgaaaatta ttaccatatg tttagagtda dtoodagsam tgtctatgaa tgttgaaaac ttttcaatst cttggratgc tcakttaatt cgcagaccgg	2005
cttaacggat taaacgcccc ccccc	•
<210> 16	
<211> 943	
<212> DNA	
<213> Homo sapiens	
<400> 16	60
changed changed detageed ecogotope	120
	180
	240
	300
	360
	420
aacaagaccc ggacatteca ggctacagte eagggtgacca cagaaaggac tgcaggaacc tctgcaggac aactggtcct tatcacagcc agggtgacca cagaaaggac tgcaggcacc catatcctgc	480
totgoaggac aactggtoot tatoataget agggtoot total total total aggetoot total aggeto	540
tgcctatatt ttagtgctgt tccaggaatt cacctagaa tgggtgagga atgtgttagt	600
tgcctatatt ttagtgctgt tctaggatte agccctagaa tgggtgagga atgtgttagt tcagaggagg gggctggaaa tgccaccctg agccctagaa tgggtgagga gctqqctctg	660
tragaggagg gggctggaaa tyttateet aggregetea ceteggagga getggetetg gtetggage atgaagget tgtgetgace aagetgetea ceteggagga gettetetge	720
	780
	840
	900
anctotant tagaaaaggt tttagaaaaa gugaayagut ggaarg ggaarg	943
aagctttttt gcccaaaaaa aaaaaaaaaa aaaaaaaaaa	
<210> 17	,
<211> 1503	
<212> DNA	
<213> Homo sapiens	
<400> 17	60
at arrange and tooctto otottgoald catgegree	120
	180
	240
	300
	360
	420
	480
	540
	600
tgggaaacag gcactgrete caggetadag coordagagat gtgctgaaat acgtccggga tgcctgtage aggaggccetg aggaggag ggaagaggat gtgccgagag cagctggagg	660
tgcctgtagc aggagccctg aggaggagga ggaggaggaggaggaggaggaggagga	
	720
	780
gatetttte agetagggea taaactgtge actgaactge oogoog getetetgtg	
gatetttte agetagggea taaactgtge actgaactge dogeosas s acagetgage ttecactggt getgetggge egecegeetg tgggaatggg getetetgtg acagetgage ttecactggt getgetgge agatteacet caggecagaa geceetggae	780
gatetttte agetaggea taaactgige actgatetge degraatgg getetetgtg acagetgage ttecactggt getgetggge egecegeetg tgggaatggg getetetgtg etectacett tgtgeettet tgggeetgge agatteacet eaggeeagaa geceetggae etectacett tgtgeettet egggeetgge agatteacet eaggeeagaa aatgagatee	780 840
gatetttte agetaggea taaactgige actgatetge of gagaatggg getetetgig acagetgage ticcactggt getgetggge egecegeetg tgggaatggg getetetggae etectacett tgtgeettet tgggeetgge agatteacet eaggeeagaa geeeetggae acteegggee ttggggetge egttetgagt gtgeggaagg eaggaeteaa aatgagatee acteegggee ttggggetge egttetgage tettgaggete tggagteeea	780 840 900
gatetttte agetaggea taaactgige actgatetge of solves acagetgage tecategge geteetgige egecegeetg tgggaatggg getetetgig etectacet tgtgeettet tgggeetgge agatteacet caggecagaa geceetggae acteeggge ttggggetge egtetgagt gtgeggaagg caggaeteaa aatgagatee catttgaete eetetgtatg taetgtgeee teteetgget ettgaggete tggagteeea catttgaete eetetgtatg taetgtgeee teteetgget ettgaggete tggagteeea	780 840 900 960 1020
gatetttte agetaggea taaactgige actgatetge of solves acagetgage tecactgig getgetgage egecegeetg tgggaatggg getetetgig etectacett tgtgeettet tgggeetgge agatteacet caggecagaa geceetggae acteetggee ttggggetge egetetgagt gtgeggaagg caggaeteaa aatgagatee cattigaete eetetgtatg taetgtgee teteetgget ettgaggete tggagteeca attgtetgtg ttagteagtg accaggitee aggaaaatg atgteatgtg gtggteeaac attgtetgtg ttagteagtg accaggitee aggaaaagga teaetgeeag gtgeaetgga	780 840 900 960 1020 `1080
gatetttte agetaggea taaactgige actgadetge oggests gggaatggg getetetgtg acagetgage ticcaetggt getgetggge egecegeetg tgggaatggg getetetggae etectacett tgtgeettet tgggeetgge agatteaeet eaggeeagaa geeeetggae acteegggee ttggggetge egttetgagt gtgeggaagg eaggaeteaa aatgagatee eatttgaete eetetgtatg taetgtgeee teteetgget ettgaggete tggagteeea ettgtetgtg ttagteagtg accaggitee agggaaaatg atgteatgtg gtggteeaae attgtetgtg ttagteagtg accaggitee aagaaaagga teaetgeeag gtgeaetgga ttaetggaae caaaagagaea gtaetttgea aagaaaagga aageetgtt caaaaaatagt	780 840 900 960 1020 `1080 1140
gatetttte agetaggea taaactgige actgatetge of solves acagetgage tecategge geteetgige egecegeetg tgggaatggg getetetgig etectacet tgtgeettet tgggeetgge agatteacet caggecagaa geceetggae acteeggge ttggggetge egtetgagt gtgeggaagg caggaeteaa aatgagatee catttgaete eetetgtatg taetgtgeee teteetgget ettgaggete tggagteeea catttgaete eetetgtatg taetgtgeee teteetgget ettgaggete tggagteeea	780 840 900 960 1020 `1080

```
ctgcccacag ttctctgccc tgtctgggag gttgaggcca cagtgtatag actggtaagc
                                                                    1260
cagacaggee tectecegea agetgetace ttgettteae etgtacettg gteeceggge
                                                                     1320
agctagctat aaagcaagag ggacaggagc ccagaagaga cactgaggac aagagatcac
                                                                     1380
                                                                     1440
accagagtac atgtctctgc ctctgttttc agtgtggctt tggacaggaa tatatgaata
                                                                     1500
aatcactgcc atacaggttt tccaatacac aagtgctaga aaatacacac aattccccaa
                                                                     1503
tga
<210> 18
<211> 1512
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (207)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (209)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (521)
<223> n equals a,t,g, or c
<400> 18
                                                                       60
geagageece tgggtgtgag aagetegtet ecegtgggtt geattggete tgeectatet
                                                                      120
ctgcctccag cacccagggc ggccgcagat ggcagtgtct ctggggacag cagctgcgaa
tgagtccacg ggccaacgct gagctgctca ggctgaggcg gtgtgctcag cacagagccc
                                                                      180
ceggaactgg catetgeagg gegtganena aggeegeege gatgeegeae tteetggaet
                                                                      240
ggttcgtgcc ggtctacttg gtcatctcgg tcctcattct ggtgggcttc ggcgcctgca
                                                                      300
                                                                      360
tctactactt cgagccgggc ctgcaggagg cgcacaagtg gcgcatgcag cgcccctgg
                                                                      420
tggaccgcsa cctccgcaag acgctaatgg tgcgcgacaa cctggccttc ggcggcccgg
                                                                      480
aggtctgagc cgacttgcaa aggggatagg csggcggcac cgggcgccct cccccagccc
                                                                      540
gcccgcccg cccagcccgg agacccccaa ggcagaggga ngccggcctg ttggccctcc
                                                                      600
acgctatece tetgeageet gggecetece gacagaggee ceaggtgege tgscagtgra
ggtggggcac ttaggtgcct ggctggccca gggcttgctc tccgtgtcaa gccgactcac
                                                                      660
                                                                      720
ccagagccca ccctcccaag ctcaggggca tcctccgctg ggccccagtg cctttgcrct
gegeageact etgeceteca etggaeteag geatgtetat ggetgeetgt eetgaggete
                                                                      780
cggagccetc atttcttcgt gaagtcccca geteceetge etecaetcaa tggcaeegge
                                                                      840
                                                                      900
cctgcaactt taggcaggtc gaagccaacc caaggaaaga acctaagaac ctcgtttgga
                                                                      960
gggatgtcag cttgggccag mccagccgca ccccgcgggg ctcaggcttg gaactggtga
                                                                     1020
1080
agagagaga agagagaga agagagagtc tggggggagc gggcaagcat ggggagatga
gatgtgtata tgtgagagag agtgtggggg ccccaggcag ggcaggaggt ggtggaaacg
                                                                     1140
                                                                     1200
 gggtgaactc cgtgggctgt gtgaggactg tccatagtgg gtccmaaccc cctccctctg
ctggagtttc ctagcccttc cccctcccya agactgwggc agcaggcagg agcccctgcc
                                                                     1260
                                                                     1320
 ctccctccct gtcctgtgcc acacttctgg ggccaaaccc agcccccttg agccaggccc
                                                                     1380
 tgccagactc caagcccacc ctagaaccct cctcctgtgt ggagactctg ttgccccact
 ttggacacag attggcaacc tgcctcaccm ckcccccctw cgctggggct tccatcttaa
                                                                      1440
 tttattctca ataataaaga cttcatgatg amaaaaaaaa aaaaaaaaaa aaaaaaaaaa
                                                                      1500
                                                                      1512
 aaaaaaaaa aa
```

```
<211> 1655
<212> DNA
<213> Homo sapiens
<400> 19
                                                                         60
ccacgcgtcc gggcaaagaa ttaaacctgg tgtttggact tcaacttagc atggctagaa
                                                                        120
ttggaagtac agtaaacatg aacctcatgg gatggctgta ttctaagatt gaagctttgt
                                                                        180
taggttctgc tggtcacaca acceteggga teacacttat gattgggggt ataacgtgta
ttctttcact aatctgtgcc ttggctcttg cctacttgga tcagagagca gagagaatcc
                                                                        240
                                                                        300
ttcataaaga acaaggaaaa acaggtgaag ttattaaatt aactgatgta aaggacttct
ccttacccct gtggcttata tttatcatct gtgtctgcta ttatgttgct gtgttccctt
                                                                        360
ttattggact tgggaaagtt ttctttacag agaaatttgg attttcttcc caggcagcaa
                                                                        420
gtgcaattaa cagtgttgta tatgtcatat cagctcccat gtccccggtg tttgggctcc
                                                                        480
tggtggataa aacagggaag aacatcatct gggttctttg cgcatagcag ccactcttgt
                                                                        540
gtcccacatg atgctggcct ttacgatgtg gaaccettgg attgctatgt gtcttctggg
                                                                        600
                                                                        660
actetectae teattgettg cetgtgeatt gtggceaatg gtggcatttg tagtteetga
                                                                        720
acatcagctg ggaactgcat atggcttcat gcagtccatt cagaatcttg ggttggccat
catttccatc attgctggta tgatactgga ttctcggggg tatttgtttt tggaagtgtt
                                                                        780
cttcattgcc tgtgtttctt tgtcactttt atctgtggtc ttactctatt ggtgaatcgt .
                                                                        840
gcccagggtg ggaacctaaa ttattctgca agacaaagga agaaataaaa tttcccatac
                                                                        900
tgaatgagaa gttaaaatga atgtgtcaga gaatgggctt aacacatcgt tggtttgaaa
                                                                        960
acttccattt taaaaattta gagtttagtc attagaaaaa ataatggact ggaaagttat
                                                                       1020
                                                                       1080
atttatatcc aaatatacct atttcaaagt gtatttgtga ggcctgtttt agcctgtgtc
                                                                       1140
ttttgtattg tgtgttgcta aagaattcta cttttagtag gctaatcaac aatgaaaggg
ttagaaaatt gctgtggaac atccaggtga acttcaggaa agacagtgaa aaatggaaaa
                                                                       1200
cgttggagct tctgttgaga taatcttcat taggtatata tcttagggat acagcctttt
                                                                       1260
ctttatctta tagcaggaaa aaaaaacttt tgagggaaat agaagggctg cgttacacaa
                                                                       1320
aataaacaat ggcattgtca taggccttcc ttttactagt agggcataat gctagggaat
                                                                       1380
atgtgaagat gtttttttga agtctctttc tgatcacgaa caatagcttg cgctctactc
                                                                       1440
                                                                       1500
tgtagttatg tggattgccg agcaatgacc cttttcaatt tcttatttct gtgttactga
ggaccctaat cacttaggga tgtaatttta tagtataaac tttctgtaca gtttttctta
                                                                       1560
                                                                       1620
tagtctaata agtaaaaagt gtccttcaaa ttatgataat tgcctatgta catggataaa
                                                                       1655
ttaaaacact gcacacggaa aaaaaaaaaa aaaaa
<210> 20
<211> 2525
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (5)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (10)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1354)
<223> n equals a,t,g, or c
<400> 20
tgacnctatn gtaaggtacg cctgcaggta ccggtccgga attcccgggt cgacccacgc
                                                                         60
gtccggtctg ccaacaaggt cgttcatgaa agtgtttttc tctttaaggt aattaaaaaa
                                                                        120
```

			,			
cagtggaatg	gaaaaacagt	gctgtagtca	tcctgtaata	tgctccttgt	caacaatgta	180
tacattcctg	ctaggtgcca	tattcattgc	tttaagctca	agtcgcatct	tactagtgaa	240
gtattctgcc	aatgaaggta	agttaagact	tggtatatgc	atggagcact	tccatctaat	300
cacacatctc	tctcttgcct	ttggttctgt	tatatataac	atggaaataa	taatgccttt	360
tgcttcatgt	gagtgataaa	gcatatttaa	atttgattat	ttaaccttgc	attcctcaac	420
aagaaaaaat	gtttgataat	ggatgaaatg	tgagtcaatc	agatacaaaa	atcaaaccct	480
ttggtgaaga	accagtcgta	acatttgact	gttaattcaa	tcaacaggtg	tttctggacc	540
tatagcaaaa	tgtgtaattg	cgccttattt	tgaagtagaa	ggatatattt	gtttggtcac	600
ttggcatttg	tgaggtactt	actattgtaa	ttattgtatc	aatggtaagg	tgtcagcatt	660
	gtcatattgt					720
aagaaatctt	tttttatgta	gggatttaag	caaacacttt	aattccacca	aactgtattg	780
agtacttctt	actagttatt	gagtgaaggg	gtgggttgcc	cctccacatc	tgtgggtgtt	840
tctcgttagg	tggaacgaga	gacttggaaa	agaaagggac	atagacaaag	tatagagaaa	900
gaaaaaaggg	ggcccagggg	accggcgctc	agcacacgga	ggatctctgc	cagcctctga	960 1020
gttccmttag	tatttattga	tcattattgg	gtgtttctcg	gagagtggga	tgtggcagga	1020
tcataggata	gtagtggaga	gagggtcaac	aggtaaacac	grgaacaaag	gtetttgeat	1140
	gtaaagratt					1200
aatgctttac	aaagcagtat	tgctgcccgc	akgtcccacc	recayeceta	aggeggeeee	1260
yccctatctc	agtagatgga	gcatacaatc	gggttttata	ccgagacacc	tteetteete	1320
	ggagacagat					1380
ttttactaat	cctcctcagc	acagacectt	actigggege	cgggccgggg	cttggacaat	1440
cetteeett	cccacgaggc cctaggcaga	catacticag	actateacat	gygyayaaac	tccctasata	1500
	gggagtggtg					1560
acaaaccaca	tcttgcaccg	cccttaatcc	arttaaccct	gagtkgacac	agcacatgtt	1620
tcagagagca	crgggttggg	ggtaaggtta	cagattgcag	aacaaaatgg	agtctcctat	1680
gtctacttct	ttctamacag	acacagtaac	aatctgatct	ctcttttccc	cacaattgag	1740
gacacataca	atcatgatat	gacctttaat	ggtctactac	ttggagagtc	agatgtgtac	1800
ccaagtctct	actgcagtta	acatttacct	gccaggcact	aggctaagta	ttagcagcag	1860
gttcaaagtg	cataagatat	agaccttgtc	ctcaagactt	agtttattag	gagagacatg	1920
aatgtaaaca	catcatgaaa	atccattata	ataactgcaa	taattgatat	atcctgaaga	1980
tgcagagatt	gtctagagga	taaagttata	tattctgttt	ggtaggggat	gatgtggagt	2040
tcaaatggat	cagagaacac	ttcgctgatt	agaagtcagt	tgatccacta	gaagtcaagg	2100
	gattcaaaac					2160
	ttgtgagaac					2220
tagccaggca	cagtggctca	tgctgtaatc	ccagcacttt	aggaggccaa	ggccggcgga	2280
tcacttgagg	tcaggaattt	gagaccaggc	tggccaacat	ggtgaaaccc	atctctacta	2340
gaaatacaaa	aaattagctg	ggtttgttgg	cgtgtgcctg	taatcccagc	tatgagggag	2400
actgaagcag	gagaatgaac	ccgggaggca	gagattgcag	tgagccgaga	tcatgccact	2460
gcactccagc	ctgggcaaca	gagcaaaact	gtctcaagaa	aaaaaaaaa	aaaaagggcg	2520
gccgc						2525
<210> 21						
<211> 1396						
<212> DNA						
<213> Homo	sapiens					
<400> 21						
	tcgcgcccgg	gaggcgccgg	agcccagcgg	ctggcgccag	atccaggctc	60
	catgtccggc					120
agctgctgtt	tgaattatct	gtgaatgttg	ggaagaggaa	tgccagagct	gccggctgaa	180
aattacccaa	ccaagagaaa	tctgcaggat	ggactttctg	gtcctcttct	tgttctacct	240
ggcttcggtg	ctgatgggtc	ttgttcttat	ctgcgtctgc	tcgaaaaccc	atagcttgaa	300
aggcctggcc	aggggaggag	cacagatatt	ttcctgtata	attccagaat	gtcttcagag	360
agccrtgcat	ggattgcttc	attacctttt	ccatacgaga	aaccacacct	tcattgtcct	420
gcacctggtc	ttgcaaggga	tggtttatac	tgagtacacc	tgggaagtat	ttggctactg	480
tcaggagetg	gagttgtcct	tocattacct	tettetacce	tatctgctgc	taggtgtaaa	540

traggagetg gagttgteet tgcattacet tettetgeec tatetgetge taggtgtaaa

WO 00/04140 PCT/US99/15849

```
600
cctgtttttt ttcaccctga cttgtggaac caatcctggc attataacaa aagcaaatga
                                                                        660
attattattt cttcatgttt atgaatttga tgaagtgatg tttccaaaga acgtgaggtg
ctctacttgt gatttaagga aaccagctcg atccaagcac tgcagtgtgt gtaactggtg
                                                                        720 .
tgtgcaccgt ttcgaccatc actgtgtttg ggtgaacaac tgcatcgggg cctggaacat
                                                                        780
caggitactic cicatctacg ictigaccti gacggccicg gcigccaccg icgccattgi
                                                                        840
gagcaccact tttctggtcc acttggtggt gatgtcagat ttataccagg agacttacat
                                                                        900
                                                                        960
cgatgacctt ggacacctcc atgttatgga cacggtcttt cttattcagt acctgttcct
                                                                       1020
gacttttcca cggattgtct tcatgctggg ctttgtcgtg gttctgagct tcctcctggg
                                                                       1080
tggctacctg ttgtttgtcc tgtatctggc ggccaccaac cagactacta acgagtggta
                                                                       1140
cagaggtgac tgggcctggt gccagcgttg tccccttgtg gcctggcctc cgtcagcaga
                                                                       1200
gccccaagtc caccggaaca ttcactccca tgggcttcgg agcaaccttc aagagatctt
tctacctgcc tttccatgtc atgagaggaa gaaacaagaa tgacaagtgt atgactgcct
                                                                       1260
ttgagctgta gttcccgttt atttacacat gtggatcctc gttttccaaa aaaaaaaaa
                                                                       1320
aaaaaaaaa aaaaaaaaa aaaactcgag ggggggcccg gtacccaatt cgccctggag
                                                                       1380
                                                                       1396
ttcaagtaga catcaa
<210> 22
<211> 1069
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (508)
<223> n equals a,t,g, or c
<400> 22
                                                                         60
ggcacgagca cagceteagg ceetgeeeca gaeetgeaga ateagaaact etggggtgag
                                                                        120
gcctggttat ctgctgtaac agaccttcca gtgggttctg atgccctcta gagcaggaga
                                                                        180
accactaget tagaggttgc agtatgtttg gcatcttgcc atttgtgtta gttcagagga
                                                                        240
atggctgacc cccatgtctc atttctaagc ttcaggcagc ttttctcctg ggcagctgtc
                                                                        300
attetgttga ggggaatect ggggaetgtg geteeteete eetgteegtg tgteettgat
                                                                        360
ctggcagtct accccttca tctccccgtg gaggctccat gcctagaggt ggtcttcaaa
cagaagaatg gcaaagataa ttgtctcgtg ttttaccctg accccattcc tttaagaggg
                                                                        420
tcacttcttg gcccattcat taaaaaccaa tgtcatagtt ctgtgattcc actatcagac
                                                                        480
agtgccacgt ccaaggcgcg ggctcttnac ctccctggaa gagagactgt gctgtctgtg
                                                                        540
cttcctgtgt tctccagtcc cacgetecca cggacecacg ccettggaga etcecteggt
                                                                        600
gtcccagggc ttctggtgtg ttcagagacc tccacactca acgaccactg gtgctgcaga
                                                                        660
                                                                        720
agggeeggtg ettacattee aattaacaga egetttteee atetaatgee tettgeette
                                                                        780
tcctaacacc acctcgggag tgtttatgtc tattctaagt gaatttcact gtgtgaaaaa
attcacacct gttgtcccag cgatttggga ggccggggcg ggtgtatcat ttgagcccag
                                                                        840
                                                                        900
gagtttgagg ctagcctggg caggatggtg aaaccccgtc tctataaaga aattttaaaa
                                                                        960
attagctggg catagtggca cgtgcctgta gttccatcta ctggggaggc tggggtggga
ggategeatg agecegggag tttgaggetg cagtgagetg tgategeage actgeactee
                                                                       1020
agtctgggca acagagcaag accctgtctc ttaaaaaaaaa aaaaaaaaa
                                                                       1069
<210> 23
<211> 1658
<212> DNA
<213> Homo sapiens
<400> 23
                                                                        60
ggcacgagcc ggcctgccag agccatgccc ctgactcctc agcttcaaaa tcaggggtct
caggacagag gatgctgggt gggctcagag ctcatcaggg gggctgtgtg tgagagggga
                                                                        120
tgccttctgg atgccctcat cctcctcggg gctggggtct ccctcaaggc cacccagctc
                                                                        180
cttcctttgt ttgctgctgc tactcctgcc gcctgctgcc ttggccctgc tgctcttctt
                                                                        240
```

PCT/US99/15849

					•
cttggacttc ttccctccca	gggcagctgt	gtctcccttc	ttgccggacc	actgctctgc	300
caggcaacct agggtgtgga	ggagagagac	cctcaacaga	agtgcctcag	ggctggggtg	360
ctgggcaagg agcactgagc	agggagccgt	gggagtagca	actgggactg	tgcttgacat	420
cagcetecet geeteetgee	tctcgctgtg	gccaccaggc	ccctctgggg	gcatctgact	480
tgtctgccca tcattctgca	cctggtttca	gtgactctta	cttcaccatg	tcttgccaat	540
caagcctttc aagggcagag	atcctacaat	gccctctggt	gccctctgtt	tctcctccta	600
cccacctccc ccaagggaga	gcaaacaaat	catccagagc	cagcctgccc	ttgcttcccc	660
aaactcactg gtgtcttttc	ccttcagcac	gtggttggcg	cagaggaatt	cagtcaggtc	720
ttcctcctgg tggatcctgt	accagtcctc	gatcacctcc	tcaaactctt	caccagcaca	780
tcacacttgt taatcataat	acctcatatt	ggcaaagccc	cagcacctga	ctcgctccta	840
gaggagctca gcctaagcct	cgcaacccac	tgcaaggtag	cagtggcacg	gttcacctaa	900
ggaaactgag gccagagagg	tgaaatgacc	tgaccaaagc	caccccggcc	tgggtggact	960
tectcagage agacecaate	cccaccagcc	cttcactggg	cacagcaacc	cttccaaggg	1020
ctgaagggcc tgtacctgct	tcttgaggtc	agccacttct	gcagaagtct	cgttcaacag	1080
ctcatagggg atgtccatca	ccaccttgac	ccctttgtgt	accaggttgt	gtaatgtctc	1140
aaaggtctct gacatgccct	ggaagaagcg	accagatatg	gcaggcggag	ctcccttctc	1200
teceteceae ectegtetee	cagtggtggc	taagaaccca	gctataagac	caatgctcaa	1260
cgccctctaa ggatcctcat	ccttttttt	ttgagaagga	gtctcactct	gtcgcccagg	1320
ttggagcgtc tcagctcact	gcaacctctg	cctcccaggt	tcaagcgatt	ttcctgcctc	1380
agcctcccaa gcagctggga	ctacaaaggc	gtgccaccat	acccggctaa	tttttgtaga	1440
gttggggttt tgtcatgttg	gtcaggctgg	tctcgaactc	ctagcatcaa	gttttccact	1500
cacctcagcc tcccaaagtg	ctgagattac	aggcgtgagc	caccgcacct	ggcctcatcc	1560
ttgacctgac cttcctcttc			cacaacccct	gcacatatac	1620
cccctgatct gcctctgcac	acctcatcgc	ttcaaaaa			1658
					•
<210> 24					
<211> 1077	,				
<212> DNA					
242					
<213> Homo sapiens					
<220>					
<220> <221> SITE					
<220> <221> SITE <222> (1036)			·		
<220> <221> SITE	or c		·		
<220> <221> SITE <222> (1036) <223> n equals a,t,g,	or c		·		
<220> <221> SITE <222> (1036) <223> n equals a,t,g,	or c				
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE	or c				
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038)					
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE					
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g,					
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g,	or c	tecttectta	cagccttaaa	tcggtctgta	
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g,	or c	tccttccttg	cagccttaaa cttattattg	tcggtctgta ttcttcctga	60 120
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgccttag	or c	ttgggtgtaa	cttattattg	ttcttcctga	
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgccttag cctacttcct gtttatcact	or c ctcccaggac aaacccacgc	ttgggtgtaa tcattttggc	cttattattg atttcggtga	ttcttcctga tcgggttgga	120
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgcttag cctacttcct gtttatcact actattgaag cccgctttca	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc	ttgggtgtaa tcattttggc cccattttcc	cttattattg atttcggtga ctttgaaagg	ttcttcctga tcgggttgga aagacttctg	120 180
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaattc cgcgcttag cctacttcct gtttatcact actattgaag cccgctttca gcttctccta aatctccgtt	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa	ttgggtgtaa tcattttggc cccattttcc ggggagtcca	cttattattg atttcggtga ctttgaaagg agcctctgtc	ttetteetga tegggttgga aagaettetg atgaggaaeg	120 180 240
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaattc cgcgcttag cctacttcct gtttatcact actattgaag cccgctttca gcttctcta aatctccgtt gaaatgcgag ggcctcgggt	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgccct	cttattattg atttcggtga ctttgaaagg agcctctgtc cagcttgcac	ttcttcctga tcgggttgga aagacttctg atgaggaacg gccggaagct	120 180 240 300
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaattc cgcgcttag cctacttcct gtttatcact actattgaag cccgctttca gcttctcta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgccct ggccttcgac	atttattattg atttcggtga ctttgaaagg agcctctgtc cagcttgcac tcgctatgtc	ttcttcctga tcgggttgga aagacttctg atgaggaacg gccggaagct cactaacaat	120 180 240 300 360
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaattc cgcgcttag cctacttcct gtttatcact actattgaag cccgctttca gcttctcta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga atgtcggacc cacggaggcc	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct gaacaaagtg	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgccct ggccttcgac ctgaggtaca	atttattattg atttcggtga ctttgaaagg agcctctgtc cagcttgcac tcgctatgtc agcccccgcc	ttcttcctga tcgggttgga aagacttctg atgaggaacg gccggaagct cactaacaat gagcgaatgt	120 180 240 300 360 420
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgccttag cctacttcct gtttatcact actattgaag cccgctttca gcttctccta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga atgtcggacc cacggaggcc aacccggcct tggacgaccc	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct gaacaaagtg gacgccggac	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgccct ggccttcgac ctgaggtaca tacatgaacc	atttcggtga ctttgaaagg agcctctgtc cagcttgcac tcgctatgtc agcccccgcc	ttcttcctga tcgggttgga aagacttctg atgaggaacg gccggaagct cactaacaat gagcgaatgt gatcttcagc	120 180 240 300 360 420 480
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgccttag cctacttcct gtttatcact actattgaag cccgctttca gcttctccta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga atgtcggacc cacggaggcc aacccggcct tggacgaccc atgtgcggcc tcatgcttaa	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct gaacaaagtg gacgccggac gctgaagtgg	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgccct ggccttcgac ctgaggtaca tacatgaacc	atttcggtga ctttgaaagg agcetctgtc cagcttgcac tcgctatgtc agccccgcc tgctgggcat tcgctgtcta	ttcttcctga tcgggttgga aagacttctg atgaggaacg gccggaagct cactaacaat gagcgaatgt gatcttcagc ctgctccttc	120 180 240 300 360 420 480 540
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgccttag cctacttcct gtttatcact actattgaag cccgctttca gcttctccta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga atgtcggacc cacggaggcc aacccggcct tggacgaccc atgtgcggcc tcatgcttaa atcagctttg ccaactctcg	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct gaacaaagtg gacgccggac gctgaagtgg	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgccct ggccttcgac ctgaggtaca tacatgaacc tgtgcttggg gacacgaagc	atttattg atttcggtga ctttgaaagg agcetctgtc cagcttgcac tcgctatgtc agcccccgcc tgctgggcat tcgctgtcta aaatgatgag	ttcttcctga tcgggttgga aagacttctg atgaggaacg gccggaagct cactaacaat gagcgaatgt gatcttcagc ctgctccttc tagcttcatg	120 180 240 300 360 420 480 540
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgccttag cctacttcct gtttatcact actattgaag cccgctttca gcttctccta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga atgtcggacc cacggaggcc aacccggcct tggacgaccc atgtgcggcc tcatgcttaa atcagctttg ccaactctcg ctgtccatct ctgccgtggt	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct gaacaaagtg gacgccggac gctgaagtgg gagctcggag	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgcct ggccttcgac ctgaggtaca tacatgaacc tgtgcttggg gacacgaagc ctgcagaatc	cttattattg atttcggtga ctttgaaagg agcctctgtc cagcttgcac tcgctatgtc agcccccgcc tgctgggcat tcgctgtcta aaatgatgag ctcagcccat	ttcttcctga tcgggttgga aagacttctg atgaggaacg gccggaagct cactaacaat gagcgaatgt gatcttcagc ctgctccttc tagcttcatg gacgcccca	120 180 240 300 360 420 480 540 600
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgccttag cctacttcct gtttatcact actattgaag cccgctttca gcttctcta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga atgtcggacc cacggaggcc aacccggcct tggacgaccc atgtgcgcc tcatgcttaa atcagctttg ccaactctcg ctgtccatct ctgccgtggt tggtgatacc agcctagaag	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct gaacaaagtg gacgcggac gctgaagtgg gagctcgag gatgtcctat ggtcacattt	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgcct ggccttcgac ctgaggtaca tacatgaacc tgtgcttggg gacacgaagc ctgcagaatc tggaccctgt	cttattattg atttcggtga ctttgaaagg agcctctgtc cagcttgcac tcgctatgtc agcccccgcc tgctgggcat tcgctgtcta aaatgatgag ctcagcccat ctatccacta	ttetteetga tegggttgga aagaettetg atgaggaaeg geeggaaget cactaacaat gagegaatgt gatetteage etgeteette tagetteatg gaegeeceea ggeetggget	120 180 240 300 360 420 480 540 600 660 720
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgcttag cctacttcct gtttatcact actattgaag cccgctttca gcttctccta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga atgtcggac cacggaggcc aacccggct tggacgaccc atgtgcgcc tcatgcttaa atcagctttg ccaactctcg ctgtccatct ctgccgtggt tggtgatacc agcctagaag ttggctgcta aacctgctgc	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct gaacaaagtg gacgccggac gctgaagtgg gagctcggag gatgtcctat ggtcacattt cttcagctgc	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgcct ggccttcgac ctgaggtaca tacatgaacc tgtgcttggg gacacgaagc ctgcagaatc tggaccctgt catcctggac	cttattattg atttcggtga ctttgaaagg agcctctgtc cagcttgcac tcgctatgtc agccccgcc tgctgggcat tcgctgtcta aaatgatgag ctcagcccat ctatccacta ttccctgaat	ttetteetga tegggttgga aagaettetg atgaggaaeg geeggaaget cactaacaat gagegaatgt gatetteage etgeteette tagetteatg gaegeeceea ggeetggget gaggeegtet	120 180 240 300 360 420 480 540 600 660 720 780
<220> <221> SITE <222> (1036) <223> n equals a,t,g, <220> <221> SITE <222> (1038) <223> n equals a,t,g, <400> 24 ggcacgaggg gaaagccatg cggaaaattc cgcgccttag cctacttcct gtttatcact actattgaag cccgctttca gcttctcta aatctccgtt gaaatgcgag ggcctcgggt gcgattcctg cagcggaaga atgtcggacc cacggaggcc aacccggcct tggacgaccc atgtgcgcc tcatgcttaa atcagctttg ccaactctcg ctgtccatct ctgccgtggt tggtgatacc agcctagaag	or c ctcccaggac aaacccacgc tccgggttca ggttcttttc ctctgggtaa gttactctaa ggcgtgatct gaacaaagtg gacgccggac gctgaagtgg gagctcggag gatgtcctat ggtcacattt cttcagctgc agggaacctg	ttgggtgtaa tcattttggc cccattttcc ggggagtcca aatccgccct ggccttcgac ctgaggtaca tacatgaacc tgtgcttggg gacacgaagc ctgcagaatc tggaccctgt catcctggac gccctttcct	cttattattg atttcggtga ctttgaaagg agcctctgtc cagcttgcac tcgctatgtc agccccgcc tgctgggcat tcgctgtcta aaatgatgag ctcagcccat ctatccacta ttccctgaat agggaacacc	ttetteetga tegggttgga aagaettetg atgaggaaeg geeggaaget cactaacaat gagegaatgt gatetteage etgeteette tagetteatg gaegeeceea ggeetggget gaggeegtet etaggettac	120 180 240 300 360 420 480 540 600 660 720 780 840

```
tattcatttg ctttctcgtt gaaacctgtt gttaataaag tttttcactc tgaaaaaaaa
                                                                   1020
aaaaaaaaa aaaaancncg agggggggcc cggaacccaa ttcsccggat agtgagt
                                                                   1077
<210> 25
<211> 1205
<212> DNA
<213> Homo sapiens
<400> 25
cccacgcgtc cgcagaggca gggcaatagt ggagttctgg cttggccaag cagcctagaa
                                                                     60
ctcaaagtcc atggccctt ctgggcctgg agaaattgga tggttatagc accaggcagc
                                                                    120
ccttgtgggt gggggacagc aaatgaggga cctctctttt ctctacactc tcctttggct
                                                                    180
cccggagatc tggcaggccc tggctggagg cataagatta gatgaggttg agctgttgga
                                                                    240
gaatgaagct gtgttgggag aagaaatgag gttgtaccgg aagatcaacg aggttgtgct
                                                                    300
                                                                    360
gtcagggaat gaggtggtac ttgggggcaa gtgaggctgc attattagat aaatgaggtt
                                                                    420
gtactgtcag gggatgaagt gtacttgtag tagagatgac gtcctgctgg atcagtcggc
                                                                    480
ttttgctcca tcagagaaca cagccacacc acaggaggaa ggagagtgtc cgactcagag
                                                                    540
gataaatgag ggtgtcctgc tggataaatg agggggcccg tcaggtgaat ggagtgctgt
tagcaaatga ggttgtactt gctggataaa tgggactggt gtgctggata aatggggttg
                                                                    600
tgctgtcagg tgaatgcatt actgctcgtg ggtgaagggc atcctgggaa taatgagggt
                                                                    660
                                                                    720
qtcctgctgg atagatgagc tgccaccacc aaatggatca gaccctgtcc atgaaggagg
                                                                    780
caccatcage aacgacgagg ttatcctgtt cccactgggg ctcctggage gtcttctggc
ccaggggaaa ctcggtgtgt gccaccctgg gttatccaag tctctctggg gagcagggtg
                                                                    840
                                                                    900
gggggctggg gagggcaggc agctgcattg tgcaccgtgg gacctctcct tcacccccaa
                                                                    960
tggatgccct actcctctcc ctggcacccc tcagtgggtc agactgcttc ggacattctc
                                                                   1020
accocactgo otgottotoa tootgootgt gtottottto tgoccagttt ggaaaagcco
ctattatgtg tcagccactc tgcccagtct tatttaatct ccctataaca cagtattact
                                                                   1080
                                                                   1140
1200
1205
aaaaa
<210> 26
<211> 1674
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1663)
<223> n equals a,t,g, or c
<400> 26
                                                                     60
cccgagcagc tgagtccctt ccctgtcttt cactcttctg gcatcggtgg ttttacttct
                                                                    120
togattgaac cotgetteet egaccecect gggaggeege ettetteagg egeeteeett
                                                                    180
ctctccacga gctcgctctg acagctgagg aactggcaag atcctgctac ccagagggtg
aatgggtatc tttcccggaa taatcctaat ttttctaagg gtgaagtttg caacggcggc
                                                                    240
cgtgattgta agcggagtaa gcaaacacct ccattgtatt agtcaccaga aaagtaccac
                                                                    300
tgtaagtcat gagatgtctg gtctgaattg gaaacccttt gtatatggcg gccttgcctc
                                                                    360
                                                                    420
tategtggct gagtttggga ctttccctgt ggaccttacc aaaacacgac ttcaggttca
aggccaaagc attgatgccc gtttcaaaga gataaaatat agagggatgt tccatgcgct
                                                                    480
                                                                    540
gtttcgcatc tgtaaagagg aaggtgtatt ggctctctat tcaggaattg ctcctgcgtt
gctaagacaa gcatcatatg gcaccattaa aattgggatt taccaaagct tgaagcgctt
                                                                    600
attcgtagaa cgtttagaag atgaaactct tttaattaat atgatctgtg gggtagtgtc
                                                                    660
aggagtgata tottocacta tagocaatoo caccgatgtt ctaaagatto gaatgcaggo
                                                                    720
tcaaggaagc ttgttccaag ggagcatgat tggaagcttt atcgatatat accaacaaga
                                                                    780
aggcaccagg ggtctgtgga ggggtgtggt tccaactgct cagcgtgctg ccatcgttgt
                                                                    840
```

```
900
aggagtagag ctaccagtct atgatattac taagaagcat ttaatattgt caggaatgat
gggcgataca attttaactc acttcgtttc cagctttaca tgtggtttgg ctggggctct
                                                                      960
ggcctccaac ccggttgatg tggttcgaac tcgcatgatg aaccagaggg caatcgtggg
                                                                     1020
                                                                     1080
acatgtggat ctctataagg gcactgttga tggtatttta aagatgtgga aacatgaggg
cttttttgca ctctataaag gattttggcc aaactggctt cggcttggac cctggaacat
                                                                     1140
                                                                     1200
cattttttt attacatacg agcagctaaa gaggcttcaa atctaagaac tgaattatat
                                                                     1260
gtgagcccag ccctgccage ctttctactc ctttgccctt ttcccgtgtt ctaatgtatt
ttgacaatgt tgtaagtgtt taccaagccg ttggtctcct aagggcctcc tgatggaaga
                                                                     1320
acagtggggt ggttcaaagt tatttctatg tttgtgttac catgttaact tttccccgag
                                                                     1380
agaaagtgtt aacattgaga ctctggcccc agattggtat cttctatgaa gatggatact
                                                                     1440
gatgggtgac attgaaaacg gcctgctttc caaatgtggt taaatgtaat tggttagccc
                                                                     1500
                                                                     1560
cagacttggg ctagagcaga aggcataggc cagggtggtt attgctatat gtgttacaga
                                                                     1620
1674
aaactcgagg gggggcccgg tacccaattc gccctatggt gantcgaatg ggct
<210> 27
<211> 1965
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (333)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1961)
<223> n equals a,t,g, or c
<400> 27
ggatcctcgc ggcggcggcg gtgcttacag cctgagaaga gcgtctcgcc cgggagcggc
                                                                       60
                                                                      120
ggcggccatc gagacccacc caaggcgcgt cccctcggc ctcccagcgc tcccaagccg
cagoggoogo goccottoag ctagotogot ogotogotot gottocotgo tgcoggotgo
                                                                      180
                                                                      240
gcatggcgtt ggcgttggcg gcgctggcgg cggtcgagcc ggcctgcggc agccggtacc
                                                                      300
agcagttgca gaatgaagaa gagtctggag aacctgaaca ggctgcaggt gatgctcctc
                                                                      360
caccttacag cagcatttct gcagagagcg cancatattt tgactacaag gatgagtctg
                                                                      420
ggtttccaaa gcccccatct tacaatgtag ctacaacact gcccagttat gatgaagcgg
                                                                      480
agaggaccaa qqctqaaqct actatccctt tggttcctgg gagagatgag gattttgtgg
                                                                      540
gtcgggatga ttttgatgat gctgaccagc tgaggatagg aaatgatggg attttcatgt
taacttttt catggcattc ctctttaact ggattgggtt tttcctgtct ttttgcctga
                                                                      600
ccacttcagc tgcaggaagg tatggggcca tttcaggatt tggtctctct ctaattaaat
                                                                      660
ggatcctgat tgtcaggttt tccacctatt tccctggata ttttgatggt cagtactggc
                                                                      720
                                                                      780
tctggtgggt gttccttgtt ttaggctttc tcctgtttct cagaggattt atcaattatg
                                                                      840
caaaagttcg gaagatgcca gaaactttct caaatctccc caggaccaga gttctcttta
tttattaaag atgttttctg gcaaaggcct tcctgcattt atgaattctc tctcaagaag
                                                                      900
                                                                      960
caagagaaca cctgcaggaa gtgaatcaag atgcagaaca cagaggaata atcacctgct
ttaaaaaaat aaagtactgt tgaaaagatc atttctctct atttgttcct aggtgtaaaa
                                                                     1020
ttttaatagt taatgcagaa ttctgtaatc attgaatcat tagtggttaa tgtttgaaaa
                                                                     1080
agctcttgca atcaagtctg tgatgtatta ataatgcctt atatattgtt tgtagtcatt
                                                                     1140
ttaagtagca tgagccatgt ccctgtagtc ggtagggggc agtcttgctt tattcatcct
                                                                     1200
ccatctcaaa atgaacttgg aattaaatat tgtaagatat gtataatgct ggccatttta
                                                                     1260
aaggggtttt ctcaaaagtt aaacttttgt tatgactgtg tttttgcaca taatccatat
                                                                     1320
ttgctgttca agttaatcta gaaatttatt caattctgta tgaacacctg gaagcaaaat
                                                                     1380
catagtgcaa aaatacattt aaggtgtggt caaaaataag totttaattg gtaaataata
                                                                     1440
                                                                     1500
agcattaatt ttttatagcc tgtattcaca attctgcggt accttattgt acctaaggga
ttctaaaggt gttgtcactg tataaaacag aaagcactag gatacaaatg aagcttaatt
                                                                     1560
```

			1.4			
actagaatgt	aattettgac	actotttota	taattagcgt	tcttcacccc	cacccccacc	1620
				ccaaagtctg		1680
						1740
				aacttttgag		
				ccagtgggtt		1800
				gctatgttaa		1860
acaacagaac	attaataaat	atctcttgtg	tagcaccttt	aaaaaaaaa	aaaaaaaaa	1920
aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	naaaa		1965
		•				
<210> 28						
<211> 1863						
<212> DNA						
<213> Homo	sapiens					
400 00						
<400> 28						
				ccgccgtcag		60
tcagattatc	ttaacaagaa	aaccaactgg	aaaaaaaat	gaaattcctt	atcttcgcat	120
ttttcggtgg	tgttcacctt	ttatccctgt	gctctgggaa	agctatatgc	aagaatggca	180
				cagctgtgga		240
				cagatcctat		300
				caagaaccta		360
				gaaagttcac		420
						480
				gatgctggag		540
				tcctccagaa		
				aagggcctca		600
				aaggacggtg		660
cgcagggggc	ggtggaagct	gccaaggttg	gggctttggc	atctctcatt	cgatccgtgg	720
cctccttctc	catctacagt	cctcacacag	gtattcagga	ataccaggat	ggcgtgccca	780
agattccaac	agcctgtatt	acggtggaag	atgcagaaat	gatgtcaaga	atggcttctc	840
				gacctaccca		900
				agaacaggtt		960
				ggatgatggc		1020
				gctgcgtcca		1080
				tggtgccttc		1140
				ggagtctgac		1200
						1260
				cagggccatc		1320
				gagccatgga		
				tctacttgat		1380
				tgtcatggat		1440
				tgttgcagac		1500
tgctgcctag	gtcctagaaa	cagtaagaaa	gaaacgtttt	catgcttctg	gccaggaatc	1560
ctgggtctgc	aactttggaa	aactcctctt	cacataacaa	tttcatccaa	ttcatcttca	1620
aagcacaact	ctatttcatg	ctttctgtta	ttatctttct	tgatactttc	caaattctct	1680
				catagaatca		1740
				aacattgttt		1800
				aaaaaaaaa		1860
	Luucuaaccc	ceggaagace	cccgaaaaaa	44444444	~~~555~55~	1863
cgc						1005
.210. 22						
<210> 29						•
<211> 1626						
<212> DNA						
<213> Homo	sapiens					
<400> 29						
cccacgcgtc	cggagccggg	ageeggtege	gggggctccq	ggctgtggga	ccgctgggcc	60
				ggctccttgc		120
				acgccgccaa		180
2-33909			J	-5 -5 - 5		

			15			
tcatattata	gagcccatgc	ctatacaaaa	acctaatata	gaagcatact	gtctacgctg	240
tgaatgcaaa	tatgaagaaa	gaagetetgt	cacaatcaag	gttaccatta	taatttatct	300
ctccattttq	ggccttctac	ttctgtacat	ggtatatctt	actctggttg	agcccatact	360
gaagaggcgc	ctctttggac	atgcacagtt	gatacagagt	gatgatgata	ttggggatca	420
ccagcctttt	gcaaatgcac	acgatgtgct	agecegetee	cgcagtcgag	ccaacgtgct	480
gaacaaggta	gaatatgcac	agcagcgctg	gaagetteaa	gtccaagagc	agcgaaagtc	540
totctttgac	cggcatgttg	tcctcagcta	attgggaatt	gaattcaagg	tgactagaaa	600
gaaacaggca	gacaactgga	aagaactgac	tagattttac	tgggtttcat	tttaatacct	660
tottoatttc	accaactgtt	gctggaaatt	caaaactgga	agcaaaaact	tgcttgattt	720
tttttttt	ttaacgtaat	aatagagaca	tttttaaaag	cacacagete	aaatcagcca	780
ataatctttt	cctattgtga	cttttactaa	taaaaataaa	tctgcctgta	aattatcttg	840
aagteettta	cctggaacaa	gcactctctt	tttcaccaca	tagttttaac	ttgactttca	900
agataatttt	cagggttttt	attattatta	ttttttattt	gtttgttttg	gtgggagagg	960
ggagggatgc	ctgggaagtg	gttaacaact	tttttcaagt	cactttacta	aacaaacttt	1020
totaaataga	ccttaccttc	tattttcgag	tttcatttat	attttgcagt	gtagccagcc	1080
tcatcaaaga	gctgacttac	tcatttgact	tttgcactga	ctgtattatc	tgggtatctg	1140
ctatatata	acttcatggt	aaacqqqatc	taaaatgcct	gataactttt	cacaaaaagc	1200
agattttctt	catgtactgt	dateteteat	gcaatgcatc	ctagaacaaa	ctggccattt	1260
agatttttt	tctaaagact	aaacataqto	traatatata	togtettact	catcttctag	1320
tagetttaag	gacaaatcct	aaacacagcc	acacttocaa	taaagaaatt	ttattttaaa	1380
cacciciaag	cctggattga	tastatata	acatttgtga	gcatttccgg	tcataataaa	1440
cccaagcctc	cctggattga	taatatatac	acatttccctta	gcactcccgg	tataaatacc	1500
aggcagctgt	ttgagctcca	atgtgtgcag	citigaacta	gggccggggc	ttaasstats	1560
tettetgaaa	ggtctaacca	ttattggata	actggctttt		ananananan	1620
	taatttttga	aacatcaaaa	aaaaaaaaa	adadadada	aaaaaaaaaa	1626
aaaaaa						1020
<210> 30						
044. 605						
<211> 605						
<211> 605 <212> DNA						
	sapiens					
<212> DNA	sapiens					
<212> DNA <213> Homo <400> 30						
<212> DNA <213> Homo <400> 30 ccacgcgtcc	gcccacgcgt	ccgggaaatg	accttggaga	ttgtagcaga	gagtgagcat	60
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc	gcccacgcgt ctgctggcca	gggccgggat	gagcccaaca	tgaaccctaa	gcttgaggac	120
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc	gcccacgcgt ctgctggcca ccgacacctc	gggccgggat cttcctgtgg	gagcccaaca tttacctccc	tgaaccctaa catacaagac	gcttgaggac catgaagttc	120 180
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg	gggccgggat cttcctgtgg gtgggccatc	gagcccaaca tttacctccc atcctcttca	tgaaccctaa catacaagac tcatcctctt	gcttgaggac catgaagttc catcctgctg	120 180 240
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat	gggccgggat cttcctgtgg gtgggccatc ctacgccttc	gagcccaaca tttacctccc atcctcttca ccgaactatg	tgaaccctaa catacaagac tcatcctctt ctgccatgaa	gcttgaggac catgaagttc catcctgctg gctggtgaag	120 180 240 300
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg	120 180 240 300 360
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg	120 180 240 300 360 420
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc	tgaacctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg	120 180 240 300 360
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac	tgaacctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga	gccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg ccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca	gccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga	gccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg ccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca	gccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg ccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca	gccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa	gccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg ccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31	gccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931 <212> DNA	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480 540
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931 <212> DNA	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgccctgta ccagctcggc agacagatgg cttccatcat	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc	tgaaccctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg	120 180 240 300 360 420 480 540 600
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931 <212> DNA <213> Homo <400> 31	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgcctgta ccagctcggc agacagatgg cttccatcat agtataaaac	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc agttggaacc	tgaacctaa catacaagac tcatcctctt ctgccatgaa tggggtccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg aaaaaaaaa	120 180 240 300 360 420 480 540 600 605
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931 <212> DNA <213> Homo <400> 31 gagagtgcct	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc	gggccggat cttcctgtgg gtgggccatc ctacgccttc ctgcctgta ccagctcggc agacagatgg cttccatcat agtataaaac	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc agttggaacc	tgaacctaa catacaagac tcatcctctt ctgccatgaa tggggtccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg aaaaaaaaa	120 180 240 300 360 420 480 540 600 605
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931 <212> DNA <213> Homo <400> 31 gagagtgcct ctgtgtaacc	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc sapiens aagcgggggt ttgggtggcc	gggccggat cttcctgtgg gtgggccatc ctacgccttc ctgcctgta ccagctcggc agacagatgg cttccatcat agtataaaac	gagcccaaca tttacctccc atcctcttca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc agttggaacc	tgaacctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgcttttttg aaaaaaaaa ccaggactgg	120 180 240 300 360 420 480 540 600 605
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931 <212> DNA <213> Homo <400> 31 gagagtgcct ctgtgtaacc gccatggcct	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc sapiens aagcgggggt ttgggtggcc ctgcaactgc	gggccgggat cttcctgtgg gtgggccatc ctacgccttc ctgcctgta ccagctcggc agacagatgg cttccatcat agtataaaac gaaagaggac cctgctgtct tcagctctgg	gagcccaaca tttacctccc atcctctca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc agttggaacc	tgaacctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg aaaaaaaaa ccaggactgg cccacatgtg acacatggtg	120 180 240 300 360 420 480 540 600 605
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931 <212> DNA <213> Homo <400> 31 gagagtgcct ctgtgtaacc gccatggcct agcctagccc	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc sapiens aagcgggggt ttgggtggcc ctgcaactgc	gggccggat cttcctgtgg gtgggccatc ctacgccttc ctgcctgta ccagctcggc agacagatgg cttccatcat agtataaaac gaaagaggac cctgctgtct tcagctctgg ggagactggg	gagcccaaca tttacctccc atcctctca ccgaactatg gaaggggccg gagctcctcc accggcccac ttccttctcc agttggaacc	tgaacctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgcttttttg aaaaaaaaa ccaggactgg cccacatgtg acacatggtg tgggtgtgtg	120 180 240 300 360 420 480 540 600 605
<212> DNA <213> Homo <400> 30 ccacgcgtcc gaggagcggc ccaaggcgcc atcctgtggc ctgttcctgg cccttcagct gactggcctg tcctgccagg ggagctctga gatcagctca aaaaa <210> 31 <211> 931 <212> DNA <213> Homo <400> 31 gagagtgcct ctgtgtaacc gccatggcct agcctagccc gtggattctc	gcccacgcgt ctgctggcca ccgacacctc ggcgtttccg ccatcttcat gaggactctc cctcctccgc gtgggcagac gatcaccca gacatatttc sapiens aagcgggggt ttgggtggcc ctgcaactgc tgggacatca cctgggcctc	gggccggat cttcctgtgg gtgggccatc ctacgccttc ctgcctgta ccagctcggc agacagatgg cttccatcat agtataaaac gaaagaggac cctgctgtct tcagctctgg ggagactggg agtgtgcca	gagcccaaca tttacctccc atcctctca ccgaactatg gaagggccg gagctcctcc accggccac ttccttctcc agttggaacc gtgttaccca ctctgggctg tccaggcctc tctgtaaagg	tgaacctaa catacaagac tcatcctctt ctgccatgaa tggggtcccc agacctccta actcccagag cccaacccaa	gcttgaggac catgaagttc catcctgctg gctggtgaag tccagcatgg ggcctgattg ttgctaacat cgctttttg aaaaaaaaa ccaggactgg cccacatgtg acacatggtg	120 180 240 300 360 420 480 540 600 605

			10			
gtccatattt	aacatotaaa	aatgtccscc	crckcatcca	cccaaacatg	ttgtacattt	420
		agcaataaca				480
		gaggccaagc				540
		caagtcttct				600
		atccctctcc				660
ttaatactgt	ccttttttt	ttttaacag	tattttatag	atttcagatg	actatgcaga	720
aacctaaaaa	accectoact	ctgggccggg	cctggggctc	cgaaattcca	aggcccagac	780
		tccagaattg				840
aacacaaaca	gacctcaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	900
		aaagggcggc				931
<210> 32						
<211> 1407						
<212> DNA	•					
<213> Homo	sapiens					
<400> 32						
gaattcggca	cgagggcagg	ctcagaagac	gatgcggggc	tgtgtgccgg	ccttcttgct	60 130
gcatgtactc	agcctcagga	gagcttgctg	cacccaggcc	gcccaggtct	tcacagcaca	120
actgcctgga	aggcaggttg	cgagaaggag	aggcggatgg	catgagcagc	aagggggacc	180 240
gatgctgtgc	agctcacacc	actccagaac	ctgacaaggc	accagcagga	ccccttgcca	300
		tgtttttgcc				360
		ctgccagatg				420
tatcccatag	gcccaggctc	tcctttcgct	tggcgtgggc	tttgetggee	cagigigige	480
tcctggctca	gcagaaacat	ccatttgagt	tggcatccct	gtagggatee	cagagegeeg	540
taagccttct	tgtgattggt	agggatggct	gragagaraga	cccaggagg	gggccaccac	600
tgccgcatct	acttctagac	tcccaaagga	geceaggete	totototo	tetetetete	660
cacgctggca	accacgagtt	tgggaagcag	cegtattete	atataattt	acaaattata	720
agtatccatg	acaggtatga	aacatattgt	antenatent	acyccacccc	acadactacg	780
tgattatctg	gaagetetaa	gatgagagca	aatgeetgat	kateataca	acactttaga	840
tactaaagec	catteringe	cgggcatgtt cctgaggtca	ggetteget	accarctra	ccaacatggg	900
aageccaage	gggtgaatta	aatacaagcc	agacataata	acacatacet	gtaatcccag	960
ctactcacca	agetagaga	ggaaaatcac	ttgaactcgg	gaggaggagg	ttgcagtgag	1020
		tccagcctgg				1080
aatacaaagc	ccattcttcc	agagtcttgt	gccttaaata	aaacacacct	ctctactata	1140
ggaagactgt	gcaatggcac	agccgcagag	cttaatttaa	gaggttgaag	tactctagga	1200
agaattcgta	gatcatcctc	agaaaagcct	taccctaata	ttctaccaga	aaaacgtctc	1260
ccaatcaccc	aggaaagctg	tccacagtag	tccccctta	tccacqqtqt	cactttccat	1320
		aaccacggtc				1380
	aaaaaaaaaa		-3	50		1407
<210> 33						•
<211> 1526						
<212> DNA						
<213> Homo	sapiens					
<400> 33						
ggcacgagaa	aaaaccttca	ggcggcccat	gggtatgccc	aagaggacag	agaacgaatg	60
cacagaaata	ttgtcagcct	tgcacagaat	ctcctgaact	ttatgattgg	ctctatcttg	120
gatttatggc	aatgcttcct	ctggttttac	attggttctt	cattgaatgg	tactcgggga	180
aaaagagttc	cagcgcactt	ttccaacaca	tcactgcatt	atttgaatgc	agcatggcag	240
ctattatcac	cttacttgtg	agtgatccag	ttggtgttct	ttatattcgt	tcatgtcgag	300
tattgatgct	ttctgactgg	tacacgatgc	tttacaaccc	aagtccagat	tacgttacca	360
cagtacactg	tactcatgaa	gccgtctacc	cactatatac	cattgtattt	atctattacg	420
cattctgctt	ggtattaatg	atgctgctcc	gacctcttct	ggtgaagaag	attgcatgtg	480

```
540
ggttagggaa atctgatcga tttaaaagta tttatgctgc actttacttc ttcccaattt
                                                                        600
taaccgtgct tcaggcagtt ggtggaggcc ttttatatta cgccttccca tacattatat
tagtgttatc tttggttact ctggctgtgt acatgtctgc ttctgaaata gagaactgct
                                                                        660
atgatettet ggteagaaag aaaagaetta ttgttetett eageeactgg ttaetteatg
                                                                        720
cctatggaat aatctccatt tccagagtgg ataaacttga gcaagatttg ccccctttgg
                                                                        780
ctttggtacc tacaccagcc cttttttact tgttcactgc aaaatttacc gaaccttcaa
                                                                        840
                                                                        900
ggatactctc agaaggagcc aatggacact gagtgtagac atgtgaaatg ccaaaaacct
                                                                        960
gagaagtgct cctaataaaa aagtaaatca atcttaacag tgtatgagaa ctattctatc
atatatggga acaagattgt cagtatatct taatgtttgg gtttgtcttt gttttgttta
                                                                       1020
tggttagact tacagacttg gaaaatgcaa aactctgtaa tactctgtta cacagggtaa
                                                                       1080
tattatctgc tacactggaa ggccgctagg aagcccttgc ttctctcaac agttcagctg
                                                                       1140
ttctttaggg caaaatcatg tttctgtgta cctagcaatg tgttcccatt ttattaagaa
                                                                       1200
aagctttaac acgtgtaatc tgcagtcctt aacagtggcg taattgtacg tacctgttgt
                                                                       1260
gtttcagttt gtttttcacc tataatgaat tgtaaaaaca aacatacttg tggggtctga
                                                                       1320
tagcaaacat agaaatgatg tatattgttt tttgttatct atttattttc atcaatacag
                                                                       1380.
tattttgatg tattgcaaaa atagataata atttatataa caggttttct gtttatagat
                                                                       1440
                                                                       1500
tggttcaaga tttgtttgga ttattgttcc tgtaaagaaa acaataataa aaagcttacc
                                                                       1526
tacataaaaa aaaaaaaaaa aaaaaa
<210> 34
<211> 1737
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1674)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1731)
<223> n equals a,t,g, or c
gtcgacccac gcgtccgccc acgcgtccgc ccacgcgtcc ggtttataaa cagaagttta
                                                                         60
aacttgtaag cttaagcttc cgtttataaa cagaagttta aaattatagg tcctgtttaa
                                                                        120
                                                                        180
cattcagctc tgttaactca ctcatctttt tgtgttttta cactttgtca agatttcttt
acatattcat caatgtctga agaagttact tatgcagatc ttcaattcca gaactccagt
                                                                        240
gagatggaaa aaatcccaga aattggcaaa tttggggaaa aagcacctcc agctccctct
                                                                        300
                                                                        360
catgtatggc gtccagcagc cttgtttctg actcttctgt gccttctgtt gctcattgga
                                                                        420
ttgggagtct tggcaagcat gtttcatgta actttgaaga tagaaatgaa aaaaatgaac
                                                                        480
aaactacaaa acatcagtga agagctccag agaaatattt ctctacaact gatgagtaac
atgaatatct ccaacaagat caggaacctc tccaccacac tgcaaacaat agccaccaaa
                                                                        540
                                                                        600
ttatgtcgtg agctatatag caaagaacaa gagcacaaat gtaagccttg tccaaggaga
tggatttggc ataaggacag ctgttatttc ctaagtgatg atgtccaaac atggcaggag
                                                                        660
agtaaaatgg cctgtgctgc tcagaatgcc agcctgttga agataaacaa caaaaatgca
                                                                        720
                                                                        780
ttggaattta taaaatccca gagtagatca tatgactatt ggctgggatt atctcctgaa
                                                                        840
gaagattcca ctcgtggtat gagagtggat aatataatca actcctctgc ctgggttata
                                                                        900
agaaacgcac ctgacttaaa taacatgtat tgtggatata taaatagact atatgttcaa
tattatcact gcacttataa acaaagaatg atatgtgaga agatggccaa tccagtgcag
                                                                        960
                                                                       1020
cttggttcta catattttag ggaggcatga ggcatcaatc aaatacattg aaggagtgta
kggggtgggg gttctaggct ataggtaaat ttaaatattt tctggttgac aattagttga
                                                                       1080
                                                                       1140
gtttgtctga agacctggga ttttatcatg cagatgaaac atccaggtag caagcttcag
                                                                       1200
agagaataga ctgtgaatgt taatgccaga gaggtataat gaagcatgtc ccacctccca
                                                                       1260
ctttccatca tggcctgaac cctggaggaa gaggaagtcc attcagatag tgtggggggc
                                                                       1320
cttcgaattt tcattttcat ttacgttctt ccccttctgg ccaagatttg ccagaggcaa
```

catcaaaaac	cagcaaattt	taattttgtc	ccacagcgtt	gctagggtgg	catggctccc	1380
catctcgggt	ccatcctata	cttccatqqq	actccctatg	gctgaaggcc	ttatgagtca	1440
aaggacttat	agccaattga	ttattctagg	ccaggtaaga	atggatatgg	acatgcattt	1500
attacctctt	aaaattatta	ttttaagtaa	aagccaataa	acaaaaacga	aaaggcaagt	1560
tacgagactg	acttatttt	aacttctgtg	tgttgagcta	ctgtaagctt	ggcttttgtt	1620
aaagacatac	agcaattagc	tatocaaaca	taagcattgt	tctgaaaaaa	aatntataga	1680
tagatatgtt	tatctcccat	aactcataac	tggggagtat	tataccccgg	nggcttt	1737
cagacacges						
<210> 35						
<211> 2242						
<212> DNA						
<213> Homo	sapiens					
	-					
<400> 35						
tcgacccacg	cgtccgggct	gccatggcgg	cggcgggccg	gctcccgagc	tcctgggccc	60
tettetegee	gctcctcgca	gggcttgcac	tactgggagt	cgggccggtc	ccagcgcggg	120
cgctgcacaa	cgtcacggcc	gagctctttg	gggccgaggc	ctggggcacc	cttgcggctt	180
tcggggacct	caactccgac	aagcagacgg	atctcttcgt	gctgcgggaa	agaaatgact	240
taatcgtctt	tttggcagac	cagaatgcac	cctattttaa	acccaaagta	aaggtatctt	300
tcaagaatca	cagtgcattg	ataacaagtg	tagtccctgg	ggattatgat	ggagattctc	360
aaatggatgt	ccttctgaca	tatcttccca	aaaattatgc	caagagtgaa	ttaggagctg	420
ttatcttctg	gggacaaaat	caaacattaq	atcctaacaa	tatgaccata	ctcaatagga	480
cttttcaaga	tgagccacta	attatggatt	tcaatggtga	tctaattcct	gatatttttg	540
gtatcacaaa	tgaatccaac	cagccacaga	tactattagg	agggaattta	tcatggcatc	600
cagcattgac	cactacaagt	aaaatgcgaa	ttccacattc	tcatgcattt	attgatctga	660
ctgaagattt	tacagcagat	ttattcctga	cgacattgaa	tgccaccact	agtaccttcc	720
agtttgaaat	atgggaaaat	ttggatggaa	acttytstgw	magtacymta	ttggaaaaac	780
ctcaaaatat	gatggtggtt	ggacagtcag	catttgcaga	ctttgatgga	gatggacaca	840
togatcattt	actgccaggc	tgtgaagata	aaaattgcca	aaagagtacc	atctacttag	900
tgagatctgg	gatgaagcag	tgggttccag	tcctacaaga	tttcagcaat	aagggcacac	960
tctagaactt	tgtgccattt	gtggatgaac	agcaaccaac	tgaaatacca	attccaatta	1020
cccttcatat	tggagactac	aatatggatg	gctatccaga	cgctctggtc	atactaaaga	1080
acacatetgg	aagcaaccag	caggeetttt	tactggagaa	cgtcccttgt	aataatgcaa	1140
actataaaga	gacacatcaa	atgtttaaag	tctactggga	gctgacagac	ctaaatcaaa	1200
ttaaggatgc	catggttgcc	accttctttq	acatttacga	agatggaatc	ttggacattg	1260
tagtgctaag	taaaggatat	acaaagaatg	attttgccat	tcatacacta	aaaaataact	1320
ttgaagcaga	tgcttatttt	gttaaagtta	ttgttcttag	tggtctgtgt	tctaatgact	1380
atactactaa	gataacaccc	tttggagtga	atcaacctgg	accttatatc	atgtatacaa	1440
ctatagatac	aaatgggtat	ctgaaaaatg	gatcagctgg	ccaactcagc	caatccgcac	1500
atttagetet	ccaactacca	tacaacgtgc	ttagtttagg	tcggagcgca	aattttcttg	1560
accetotota	cgttggtatt	ccccatccat	ctggagaaaa	atctatacga	aaacaagagt	1620
ggactgcaat	cattccaaat	teccagetaa	ttatcattcc	ataccctcac	aatgtccctc	1680
ggactgcaat	tacceaaac	tatcttacac	caagtaatat	tattctactt	actgctatag	1740
statestag	tatatatat	ttcatcttqq	caataattaa	cattttacat	tggcaggaaa	1800
ccccaccgg	tastagaga	eascracear	aacccacca	grttcatttt	gatgctatgt	1860
agaaaycaga	taatattana	taatoraato	actattcact	tgattagttg	aaacacaaat	1920
tetacetta	aaaaataaaa	cacattaaat	attatttata	aatgatgtat	cccatggtaa	1980
ttattaass	atattasst	aaatataatt	tgaatatgtg	acaaggtctt	ttttttaaa	2040
ccactygada	tataasaatt	tagattetet	attctgtagt	gctgtacatt	tttgttcctt	2100
tataceetat	attacetate	ctccactatt	tatatattta	taatcttatt	tgcatcatga	
tastassass	anttetatas	ataaaaataa	ttaaatgagg	aggaaaaaaa	aaaaaaaaa	
						2242
aaaaddddd	aagggcggcc	Ac				

<210> 36 <211> 2235

<212> DNA

<213> Homo sapiens

<400> 36		
gtaattcggc acgagggttc caccaacatg gagctctcgc agatgtcgs	sa gctcatgggg	60
ctgtcggtgt tgcttgggct gctggccctg atggcgacgg cggcggta	sc gcgggggtgg	120
ctgcgcgcgg gggaggagag gagcggccgg cccgcctgcc aaaaagcaa	aa tggatttcca	180
cctgacaaat cttcgggatc caagaagcag aaacaatatc agcggatt	cg gaaggagaag	240
cctcaacaac acaacttcac ccaccgcctc ctggctgcag ctctgaag	ag ccacagcggg	300
aacatatctt gcatggactt tagcagcaat ggcaaatacc tggctacc	tg tgcagatgat	360
cgcaccatcc gcatctggag caccaaggac ttcctgcagc gagagcac	cg cagcatgaga	420
gccaacgtgg agctggacca cgccaccctg gtgcgcttca gccctgac	tg cagagccttc	480
atcgtctggc tggccaacgg ggacaccctc cgtgtcttca agatgacc	aa gcgggaggat	540.
gggggctaca ccttcacagc caccccagag gacttcccta aaaagcac	aa ggcgcctgtc	600
atcgacattg gcattgctaa cacagggaag tttatcatga ctgcctcc	ag tgacaccact	660
gtcctcatct ggagcctgaa gggtcaagtg ctgtctacca tcaacacc	aa ccagatgaac	720
aacacacacg ctgctgtatc tccctgtggc agatttgtag cctcgtgt	gg cttcacccca	780
gatgtgaagg tttgggaagt ctgctttgga aagaaggggg agttccag	ga ggtggtgcga	840
gccttcgaac taaagggcca ctccgcggct gtgcactcgt ttgctttc	tc caacgactca	900
cggaggatgg cttctgtctc caaggatggt acatggaaac tgtgggac	ac aratgtggaa	960
tacaagaaga agcaggaccc ctacttgctg aagacaggcc gctttgaa	ga ggcggcgggt	1020
geegmgeegt geegeetgge eeteteece aaegeecagg tettggee	tt ggccagtggc	1080
agtagtattc atctctacaa tacccggcgg ggcgagaagg aggagtgc	tt tgagcgggtc	1140
catggcgagt gtatcgccaa cttgtccttt gacatcactg gccgcttt	ct ggcctcctgt	1200
ggggaccggg cggtgcggct gtttcacaac actcctggcc accgagcc	at ggtggaggag	1260
atgcagggcc acctgaagcg ggcctccaac gagagcaccc gccagagg	ct gcagcagcag	1320
ctgacccagg cccaagagac cctgaagagc ctgggtgccc tgaagaag	tg actctgggag	1380
ggcccggcgc agaggattga ggaggaggga tctggcctcc tcatggca	ct gctgccatct	1440
ttcctcccag gtggaagcct ttcagaagga gtctcctggt tttyttac	tg gtggccctgc	1500
ttcttcccat tgaaactact cttgtctact taggtctctc tcttcttg	ct ggctgtgact	1560
cctccctgac tagtggccaa ggtgcttttc ttcctcccag gcccagtg	gg tggaatctgt	1620
ccccacctgg cactgaggag aatggtagag aggagaggag	ag aatgtgattt	1680
ttggccttgt ggcagcacat cctcacaccc aaagaagttt gtaaatgt	tc cagaacaacc	1740
tagagaacac ctgagtacta agcagcagtt ttgcaaggat gggagact	gg gatagcttcc	1800
catcacagaa ctgtgttcca tcaaaaagac actaagggat ttccttct	gg gcctcagttc	1860
tatttgtaag atggagaata atcctctctg tgaactcctt gcaaagat	ga tatgaggcta	1920
agagaatatc aagtccccag gtctggaaga aaagtagaaa agagtagt	ac tattgtccaa	1980
tgtcatgaaa gtggtaaaag tgggaaccag tgtgctttga aaccaaat	ta gaaacacatt	2040
ccttgggaag gcaaagtttt ctgggacttg atcatacatt ttatatgg	tt gggacttctc	2100
tcttcgggag atgatatctt gtttaaggag acctcttttc agttcatc	aa gttcatcaga	2160
tatttgagtg cccactctgt gcccaaataa atatgagctg gggattaa	aa aaaaaaaaaa	2220
aaaaaaaaa ctcga		2235

<210> 37

<211> 2971

<212> DNA

<213> Homo sapiens

<400> 37

· = 0 0 - 5 .						
gacgtgagga	gcgttccatt	tggccagtgg	tgggcggttg	ccacagctgg	tttagggccc	60
		caggaggaga				120
		gtgattccct				180
		gttcaggccg				240
taaaccaatt	tatcanaga	agattgtctt	ccagggctag	caattggact	tttgatgatg	300
tttgacccac	caccaggagg	agcaggcaac	grgatttcaa	agctgggctc	agcctctgtt	360
tettetete	tataataaca	aaacccattt	tagaacaaaa	attccaatca	tatctatat	420
	cycaaccyca	cacggaaatg	cggageagga	ccaggcagga	acaccttttg	480
ggtggtgaga	aagaaggtga	tggcccggca	ggagaaact	trctacctga	cccttttcct	540
CEGEGAEGGC	cacatata	Laucecadea	aaayyycatt	ccccccc		3.0

WO 00/04140

catcctgggg	acatgtacac	tcttcttcgc	ctttgagtgc	cgctacctgg	ctgttcagct	600
gtctcctgcc	atccctgtat	ttgctgccat	gctcttcctt	ttctccatgg	ctacactgtt	660
gaggaccagc	ttcagtgacc	ctggagtgat	tcctcgggcg	ctaccagatg	aagcagcttt	720
	gagatagaag					780
tcgtatcaag	aatttccaga	taaacaacca	gattgtgaaa	ctgaaatact	gttacacatg	840
	cggcctcccc					900
cttcgaccat	cactgcccct	gggtggggaa	ttgtgttgga	aagaggaact	accgctactt	960
ctacctcttc	atcctttctc	tctccctcct	cacaatctat	gtcttcgcct	tcaacatcgt	1020
ctatgtggcc	ctcaaatctt	tgaaaattgg	cttcttggag	acattgaaag	aaactcctgg	1080
aactgttcta	gaagtcctca	tttgcttctt	tacactctgg	tccgtcgtgg	gactgactgg	1140
	ttcctcgtgg					1200
	aatcgcgtcc					1260
	tgtggcccct					1320
	ggaagtcgac					1380
	cccacagaac					1440
	ccacctccag					1500
	tatggaagag					1560
	aaacctgaga					1620
	cacccagttg					1680
	cagtggcaga					1740
	cactggttct					1800
tccctccaga	tcccagccct	cctgcttggg	gtcactggtc	tcattctggg	gctaaaagtt	1860
	gctcaaatcc					1920
	gccaggggat					1980
	tcagaagatt					2040
	acaggtacct					2100
	ctgacaccta					2160
	ataattgaca					2220 2280
	aagaatttat					2340
	ttctcattag					
	aatggcagta					2400 2460
	ctcacttaaa					2520
cctgaatgca	caaatgggaa	accaaggcac	agagaggete	teeteteete	ctcacctcac	2520
cgatgtaccc	tcaaaaaaaa	aaaaaatgct	aaccagttet	cost	teggeegag	2640
tgagggaaag	cccagcactg	ctgecetete	gggtaactca	castattasa	gracttorac	2700
	gtaaccacac					2760
	ccagagccac					2820
	ctctatttca					2880
	gttttcctct					2940
	ccggaacttt			adactgcgcc	taatgtaaaa	2971
aaaaaaaaa	aaaaaaaaa	aaaaaaaaaa	4			23.2
<210> 38						
<211> 1163						
<212> DNA						
<213> Homo	sapiens					
<400> 38						
ccacgcgtcc	gccaagggtg	ctgattaggg	aatggttatg	gactaggagt	atcagtaaca	60
atggttagaa	agtggctaac	atttgttgag	cacctgctgt	gtgcctggcc	ccggctggga	120
gccttcgtgc	ccagagtgac	cccgtctgaa	tgcagttctt	tgcctcattc	aaactgggga	180
gtgggaggca	gagctgcaca	actcacaggt	gccgagctca	agactcactc	ctgggtctgc	240
ctgggctggg	ctgtgcttgt	tgcccctgtg	gccaacacac	gcgcaccttt	cacctgaaag	300
ccaggatccg	cagaacgttc	cccgaggagg	tcattgtttg	gcactatgat	ttgtctcttc	360
ctaaaaaggt	gatagagtta	cactggagag	agcagcatcc	aggtgcagca	gggatgggcc	420
tggggctcac	gggcaggggc	tctgtgtccg	gctggggcct	ggggtcctgc	gctgcacctg	480
	gcactcagta					540

tgtgtcagaa gcactcagta aatctttgct gatgaaggat gacaggatat aggacatgat

PCT/US99/15849

1932

21

```
600
gettgetget geattgeetg caateetgga tgaatgeeca ggttggettt geteecegte
                                                                     660
gggtggatgt gacgttagct gtgatgttag gtccctggct ttaaaatacg acggaactgg
                                                                     720
gaattgaggg agcagttggg gcagaaagga cagccccgca gaggcctgga gctgagcagt
gegggegace caggageagt gagtgettee gteacageet teategeace etgtggteet
                                                                     780
cataaagggg atggaatcta cgaatttagt tttcccagcc tccttaaaaa ctcattcatg
                                                                     840
ccaggggcag tggctcacac ctgaaatccc accactttgg gaggctgagg caggctgatt
                                                                     900
acttgaggtc aggagtttga gaccagccta gccaacatgg tgaaaccccg tgtctactca
                                                                     960
                                                                    1020
aagtacaaaa aaaaaaatta gtcagacgtg gtgtcacgca cctgtaatcc cagctctttg
                                                                    1080
ggaggctgag gcaggagaat cacttgaacc caggaggcag aggttatagt gagccagtat
                                                                    1140
1163
aaaaaaaaaa aaaaaaaaaa aaa
<210> 39
<211> 1932
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1624)
<223> n equals a,t,g, or c
<400> 39
ggcacgagcc aggcccctgg gccgggcgct gaggcgggcc cctctgggca gggcccgggc
                                                                      60
ggggctgggt gggccgcccc tgctgctgcc gtccatgctg atgtttgcgg tgatcgtggc
                                                                     120
                                                                     180
ctccagcggg ctgctgctca tgatcgagcg gggcatcctg gccgaratga agcccctgcc
cctgcacccg cccggccgcg arggcacagc ctggcgcggg aaagccccca agcctggggg
                                                                     240
cctgtccctc agggctgggg acgcggactt gcaagtgcgg caggacgtcc ggaacaggac
                                                                     300
cctgcgggcg gtgtgcggac agccaggcat gccccgggac ccctgggact tgccggtggg
                                                                     360
                                                                     420
gcagcggcgc accctgctgc gccamatcct cgtaagtgac cgttaccgct tcctctactg
                                                                     480
ctacgtcccc aaggtggcct gctctaactg gaagcgggtg atgaaggtgc tggcaggcgt
cctggacagc gtggacgtcc gcctcaagat ggaccaccgc agtgacctgg tgttcctggc
                                                                     540
                                                                     600.
cgacctgcgg cctgaggaga ttcgctaccg cctgcagcac tactttaagt tcctgtttgt
gegggageee ttggaaegee teetetetge etacegeaae aagtttggeg agateegaga
                                                                     660
gtaccagcaa cgctatgggg ctgagatagt gaggcggtac agggctggag cggggcccag
                                                                     720
                                                                     780
ccctgcaggc gacgatgtca cattccccga gttcctgaga tacctggtgg atgaggaccc
                                                                     840
tgagcgcatg aatgagcatt ggatgcccgt gtaccacctg tgccagcctt gtgccgtgca
ctatgacttt gtgggctcct atgagaggct ggaggctgat gcaaatcagg tgctggagtg
                                                                     900
ggtacgggca ccacctcacg tccgatttcc agctcgccag gcctggtacc ggccagccag
                                                                     960
ccccgaaagc ctgcattacc acttgtgcag tgccccccgg gccctgctgc aggatgtgct
                                                                    1020
gcctaagtat atcctggayt tytccctctt tgcctaccca ctgcctaatg tcaccaagga
                                                                    1080
                                                                    1140
ggcgtgtcag cagtgaccat gggtgtgggg ccagcagctg gtggggactg gtttcaacgc
cagetttetg tgettetgee tgteattegg agaaactetg getetgggge ttggggette
                                                                    1200
                                                                    1260
tcaggatcct ggatggcaga gactgccctc agaarttcct tgtccagggt gggcacccac
                                                                    1320
agtgactcag aggacagggc taggcaggag acctgctgct cctcattggg gggatctctt
                                                                    1380
ggggggcaga caccagtttg ccaatgaagc aacacatctg atctaaagac tggctccaga
ccccgggctg ccaggattat gcagtccact tggtctacct taatttaacc tgtggccaaa
                                                                    1440
ctcagagatg gtaccagcca ggggcaagca tgaccagagc cagggaccct gtggctctga
                                                                    1500
                                                                    1560
tcccccattt atccacccca tgtgcctcag gactagagtg agcaatcata ccttataaat
gacttttgtg cctttctgct ccagtctcaa aatttcctac acctgccagt tctttacatt
                                                                    1620
tttnccaagg aaaggaaaac ggaagcaggg ttcttgcctg gtagctccag gacccagctc
                                                                    1680
tgcaggcacc caaagaccct ctgtgcccag cctcttcctt gagttctcgg aacctcctcc
                                                                    1740
ctaattctcc cttccttccc cacaaggcmt ttgaggttgt gactgtggct ggtatatctg
                                                                    1800
gctgccattt ttctgatgca tttatttaaa atttgtactt tttgatagaa cccttgtaag
                                                                    1860
1920
```

aaaaaaaaa aa

PCT/US99/15849

```
<210> 40
<211> 881
<212> DNA
<213> Homo sapiens
<400> 40
                                                                         60
gaattcggca cgagggaacc cagaagatgc tgcctctcct gatcatctgt ctcctgcctg
ccattgaagg gaagaactgc ctccgctgct ggccagaact gtctgccttg atagactatg
                                                                        120
acctgcagat cctctgggtg accccagggc cacccacaga actttctcaa agtattcact
                                                                        180
ccttgttcct agaggataat aattttctca aaccctggta ccttgatcgt gaccatttgg
                                                                        240
aagaagaaac agccaaattc ttcactcaag tacaccaagc cattaaaacg ttacgagatg
                                                                        300
                                                                        360
ataaaacagt acttctggaa gagatctaca cgcacaagaa tctctttact gagaggctga
                                                                        420
ataagatatc tgatgggctg aaggagaagg gagccccacc cytctccatg aatgccttcc
                                                                        480
cggctccatc tcctacttgc accccagaac cccttggctc tgtctgcctc cccagcacct
cagtttctct accttctcac ctccctggca gcctgcaatg agtcctgtgc caggaaccgg
                                                                        540
cggacctccc tgtgggctgt gagtctcagc agtgctctac tcctggccat agctggagat
                                                                        600
gtttctttta ctggcaaagg aagaaggagg cagtaaagga acagggcagc ccgcatgtct
                                                                        660
tccagaagtg aacagaggcc gcagctacca ccgtcacaaa gttcactcat ctctgggtcc
                                                                        720
cggtgacccc atcccccat accctccatc ctgggtcctg gggccccaaa gctctgaggc
                                                                        780
ctaggagact gcgctgtctc gtggtttgcc tactcctaca cctttgtaaa gagtctcttc
                                                                        840
                                                                        881
attaaaaccc ctcttcataa aaaaaaaaaa aaaaaactcg a
<210> 41
<211> 1932
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (2)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1022)
<223> n equals a,t,g, or c
<400> 41
cncggcgcgg ctcggctcat gcccccgggc gcggggcaca caggccggcc ggcagccgct
                                                                         60
gggaaatagg cccccggggg cggtggcggc ggcggggcca tggcgcggag accccgggcg
                                                                        120
ccggccgcct ccggggagga gttctccttc gtcagcccgc tggtgaaata cctgctcttc
                                                                        180
ttcttcaaca tgctcttctg ggtgatttcc atggtgatgg tggctgtggg tgtctacgct
                                                                        240
cggctaatga agcatgcaga agcagcccta gcctgcctgg cagtggaccc tgccatcctg
                                                                        300
ctgatcgtgg tgggtgtcct catgttcctg ctcaccttct gtggctgcat tgggtccctc
                                                                        360
cgcgagaaca tctgcctcct gcagacgttc tccctctgcc tcaccgctgt gttcctgctg
                                                                        420
cagctggccg ctgggatcct gggcttcgtc ttctcagaca aggctcgagg gaaagtgagt
                                                                        480
                                                                        540
gagatcatca acaatgccat tgtgcactac cgagatgact tggatctgca gaacctcatt
                                                                        600
gattttggcc agaaaaagtt tagctgctgt ggagggattt cctacaagga ctggtctcag
aacatgtatt tcaactgctc agaagacaac cccagtcgag agcgctgctc tgtgccttac
                                                                        660
tcctgttgct tgcctactcc tgaccaggca gtgatcaaca ctatgtgtgg ccaaggtatg
                                                                        720
caggcctttg actacttgga agctagcaaa gtcatctaca ccaatggctg tattgacaag
                                                                        780
ttggtcaact ggatacacag caacctattc ttacttggtg gtgtggctct aggcctggcc
                                                                        840
atcccccagc tggtgggaat tctgctgtcc cagatcctag tgaatcagat caaagatcag
                                                                        900
atcaagctac agctctacaa ccagcagcac cgggctgacc catggtactg agaatccatc
                                                                        960
ctgcacctcc tcaccatgga aactggcaag cctcataaac gaacagcagt gggtgctgaa
                                                                       1020
ancagcacca aatggagatt tggattccag cccccagtg acagcccagt gggaagaagc
                                                                       1080
```

```
aaactccaga tgggcagaag gcagggtgca caggtggctc cagtctcagg aggatgcgcc
                                                                     1140
                                                                     1200
tcctctcccc catcccagcc ctcagcattg tgccagagtg atacccttaa gtgtttgggt
ttatgttttc agttttgttt gggaaacagc agttgcacag agagttgggg gtactgctgc
                                                                     1260
tgccttttca ccgaggcact gccaccacca gctctascag ggatgctcct gagcttggcg
                                                                     1320
                                                                     1380
gacatactta gatcctaacg tgccagtgag acctggctgt ggagagtagc actggcagcc
ctgcctggac tccacttggc atgataccag ctccagaagg gaagggagtg gagcaggcag
                                                                     1440
tgaggagaga gcctggggt cggctgggga cagccgtatg tgctaggtag gagtggaggg
                                                                     1500
                                                                     1560
agatatgttt accaaatgcc tgtcctgcca tcctcccagg tagtcagagt gagctacatc
                                                                     1620
ctgccccgcc ttcatttcca tggaaacatg gcagctagga cacggggtac aacagcagcc
                                                                     1680
aaattottoo coacotooot tacttogaaa aaaagtttgg aaccotggto cotatactot
gcagtcagaa gtgggactga gccatacatg cccttgaatt cctccctgtc tggccctccc
                                                                     1740
tctccagcaa gcagggtttt ctttaacttg gcagtgtgca gaggagaagt ggtaacaccc
                                                                     1800
                                                                     1860
ccaccccatt cccctgcatc ggagctcagt attcctacag ggtaagaggt aggaatcttg
1920
                                                                     1932 -
aagggcggcc gc
<210> 42
<211> 1164
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (582)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (592)
<223> n equals a,t,g, or c
<400> 42
                                                                       60
ggcacgagct tgtgtgtcac cagcctcctg atctgccagg gtctgctctg ggttggcact
                                                                       120
gaccagggtg tcatcgtcct gctgcccgtg cctcggctgg aaggcatccc caagatcaca
                                                                       180
gggaaaggca tggtctcact caatgggcac tgtgggcctg tggccttcct ggctgtggct
                                                                       240
accagcatcc tggcccctga catcctgcgg agtgaccagg aggaggctga ggggccccgg
                                                                       300
gctgaggagg acaagccaga cgggcaggca cacgagccca tgcccgacag ccacgtgggc
cgagagctga cccgcaagaa gggcatcctc ttgcagtacc gcctgcgctc caccgcacac
                                                                       360
ctcccgggcc cgctgctctc catgcgggag ccggcgcctg ctgatggcgc agctttggag
                                                                       420
cacagegagg aggaeggete catttacgag atggeegaeg acceegaegt etgggtgege
                                                                       480
agccggccct gcgcccgcga cgcccaccgc aaggagattt gctctgtggc catcatctcg
                                                                       540
 gcgggcaggg ctaccgcaac tttggcagcg ctctgggcac antgggaagc angccccgtg
                                                                       600
 tggggagacg gacagcaccc tcctcatctg gcagtgccct tgatgctata gcgcctcccc
                                                                       660
                                                                       720
 tetecetea gagggeacag etgeaggeet gaccaaggee aegeeegget etegtgetet
                                                                       780
 aggacctgca cgggacttgt ggatgggcct ggactctcca gaaactactt gggccagagc
                                                                       840
 aaaggaaaac ctcttgtttt aaaaaaattt ttttcagagt gttttgggga ggagttttag
                                                                       900
 ggcttgggga gagggaggac acatctggag gaaatggcct tctttttaaa agcaaaaaac
 acaaaacctc acaactgcct ggcaagccca gtatcacttg tttgggccct agcgggactc
                                                                       960
 caaggcagcc acacgcccct cctggaaggg tgtgtgcgtg tgagtgtgtg cgagtgtgtg
                                                                      1020
                                                                      1080
 ggctggtgtg tgaatatcta taaataagta tatatggtgt atattatatg tgtataaata
                                                                      1140
 aagtctgtac atattggagc tctgggagat gctggaataa aagacaagag ttacatctgg
                                                                      1164
acttggaaaa aaaaaaaaaa aaaa
```

<210> 43

<211> 1105

<212> DNA

780

840

900

960

1020 1080

1140

1200 1260

1262

<213> Homo sapiens

<400> 43			•			60
gaattcggca	cgagaacaaa	ttgaaaccat	ctggtcatga	acttttattt	gttaagaggt	
tttctaatat	tgattcaatc	tctttgctta	ttctaaatgt	gttcagattt	tccacttctt	120
gagtcaattt	ggtaatttat	ttgtttctag	gaatttgtcc	atttcatcta	gtttacctaa	180
tttttgacat	ataaaattat	atatggaaat	ttctaaaata	tttaaaaatt	tctgtaatgt	240
caatagtaat	gtcccctctt	ttgttaccaa	tttgttattt	gaatcttctc	ctttttttg	300
tcaatctagc	taaaaatttg	tcaattttgt	tcgtctcttc	aaaaaaatat	acgtttgtct	360
tcatgatttc	tctawtgttt	ttccatcyat	atttcatttg	aatacatttt	taaacyttay	420
ctttattatt	tcattccttc	tgggagcttt	gggtctcatt	ttttttcct	gataatctag	480
ttatttatta	tataagatta	agtatttatt	tgaaatctgt	atgttcttta	atgtaggcat	540
tcactactat	aaatttactt	ctcaggagca	tctctgccgc	attccatgtt	ttagtatgtt	600
gtgttttaat	ttgtattcat	aactagaggg	aaacagaggt	gacggagaaa	aagacgtaca	660
aatatcatcc	acttgcaaag	tatagatttg	tttgtattgk	ratatgaatr	aaaatattac	720
gagacagata	agaaaatttg	aacactgacc	attgatgcag	ttacagttaa	ttttaaaatc	780
aaggttaata	acattttagt	tattttaaag	aatgatagta	atttagagat	gtattctgaa	840
tottttaaa	tgaaaagata	tgcctgggat	ttcttccaaa	atgaatcttg	taggttggga	900
agaaaatgag	aacatagtgg	aaacaagact	gacaatgagt	tgttgaggtt	gggcaatgcg	960 [.]
tacactaaag	cttattttat	cttattttac	tgtatatact	gttaaagctt	gcattatttt	1020
cataaatgca	tttgctaagt	gcaactgtta	tcaaataaag	tggattgggc	tctaaaaaaa	1080
	aaactcgagg					1105
<210> 44						
<211> 1262						
<212> DNA						
<213> Homo	sapiens					•
	•					
<400> 44						
cagcatgtac	ccagttgttc	tttctcctga	gaaagcaaaa	tgcctgatat.	ttcttataat	60
ccaggctgcc	acgtttacct	tgtaaaatca	atacttaatt	tttagatttt	tatattatct	120
tttctcgtga	agcaagactt	ctaaattatg	gctataatat	cttttgaatt	gttgttctta	180
atgaatette	caactgtaaa	ctcatctaat	ttcaaactta	tcatacctga	ggatgtaaca	240
ttatccttta	tttctcatct	tgatattacc	gtcaatcatt	ttgtatttct	gagtacattt	300
gaacttgctg	gagtaataga	gggaaaacct	ctgcctgatt	ctaaatcaga	tctttgtcct	360
atactcggac	aattatoott	tcatatttta	ttattttta	ttttctgggt	ttaacaaatg	420
agataacatt	ttagacataa	tatttgtaaa	catcttgact	tatttcagca	ttttcctttt	480
ttgtgtatct	tcagagagtt	tgttgaaagt	agcaatttcc	aagtaatttt	aaattattga	540
agtctactag	cacgaaaggt	caaattctta	ggatatttaa	aaaatgttgt	ttaataatca	600
aactcatctt	aaaaaatgtt	catcagactc	tgtctttgat	gcacattttg	ccaaaagaga	660
		atacantato				720

gccttatttc tgtgaaagaa atacagtatg tactttggga tttactaaag taaaactgtt

actttaaggc acagagcaga tatagaatcc ccctctctcc ccactcctag tgactggtat

tctacattaa tatttatctt ccatgcatag tgtacttgag ggaaaaaaac aataactctt

aattgtttaa tatcaaacaa taaaatcctg tgtatcagtg actgtcaata gatggctttc tgtttaaaaa ctgaagctac tccagaagta ggaattaatt tatttagtaa acaaagtcag

tcaaaccaga gccatgtcct ggggaactgt caaaagaatg gttcctaagg gccagaggcc

acatccactg gtagatgaca gaacaaccat acttcagatg gcaaaaccgg tcagtttggt

ttgcgttgtg tgcctatcct ctttctgtgt gcttcagctg aattaagtgc ttggagagct caaatagttc aagatagcca agatgaccaa ttctgccagg tggcaagcct gatcttgcaa

<210> 45 <211> 517 <212> DNA <213> Homo sapiens

```
<400> 45
gaattcggca cgagtgcact tccaccagct atgtatgaga cttcccattg ctccacatct
                                                                         60
ccagtatttt atgtggtcag tccttttgtt tttggtcatt ttggtggata tgaaatggca
                                                                        120
totcagtgtg gottttcatt atatttcctt gatgactaat ggtattcttt caccetttca
                                                                        180
gtgcttattg gccattcatg tatctttgtt ttttgtgtag cacttcaggt cttttgccca
                                                                        240
                                                                        300
tagatttagt gggttgattg ctctttatta atgatttgta gggatgttat atatattctg
                                                                        360.
gacacaagat tattgttaga gatacgtact tcagatattt tctcccagtc tgtagcttgc
ctaattatta ttattattat tatttgagat gaagtctcac tctgtcgccc aggcagaggt
                                                                        420
tgcagtgggc cgagatagca ccactacact caagcctggc tgacagagtg agactctgtc
                                                                        480
                                                                        517
tcaaaaaaaa aaaaaaaaaa aaaaaaaaaa aactcga
<210> 46
<211> 858
<212> DNA
<213> Homo sapiens
<400> 46
agaaaaaatc ctacatggat attggtagga aagagagaaa ggaagtggcc agtgtcccgt
                                                                          60
ggcctcttcc accttctgga ttgttgaagc tggggcctgg aggggatggt cctgccactc
                                                                         120
agcagggggc actaatggga ccaagctaac ctgtccagtg agaatcctgc agggagacct
                                                                         180
gagggtacca ggaaagtgca ggggaaggcc cgggaaatgg agagagctgg tctggagggg
                                                                         240
                                                                         300
aggagcaagc cgcgtggggc aggccatgtg ccttttgcct gggggagtat tactcatttg
gagetgtgcg tetggaacge etgeeteaca cacaaaggae tggggcagat gtaagttete
                                                                         360
tgcagcaacg aagcgcacag ctgagagtaa cttagaaagc acccagctaa tgctggcatc
                                                                         420
ccagatcgac cccctcctcg ctgaatgttg gcatctctgt gcctcagttt cctcatctgt
                                                                         480
aaatgggggt gataagaaat gtgtacacac ctcccgggca gtggggagga ttaaactgtg
                                                                         540
ctctgacacg atccgggcat gttcagggtg gtatctgcag taaaccgcgc tcggaaaatg
                                                                         600
gcggcgcatc agggccagcg gtgggagctc tccgtgcttg gcttgacgcc attgtggagg
                                                                         660
tggaggaggg gctgcaagac tctgagcagg aagaccccgc aaagcaggaa agcagagcca
                                                                         720
gagttggggg ccagccgcag aaacgagagc ccccgtgact ttgaggcacc ctttggagag
                                                                         780
ggcaggaagc aggaagggta aattttctcc aaaacccaag aggcagagtg accccacatg
                                                                         840
                                                                         858
ataactgagt ttctcgag
<210> 47
<211> 6107
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (5749)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (5892)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (5896)
<223> n equals a,t,g, or c
<220>
<221> SITE
```

```
<222> (5906)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (5957)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (5966)
<223> n equals a,t,g, or c
<400> 47
gcagttagtt ccttgatgtc agtagtgggc taaaggcagc ttactgtgtg tttgctggag
ctttcactca gccaagtgtt agagtcagga aacccattga ggcaatggcg tcaaatggtg
                                                                      120
                                                                      180
tttcacaaga atgagccatt cagtctttgc tcactatata tttaatattt tattattgtt
gttattgtta ttattaattg gctttctgta ttctatgcct tttatttata aagacactaa
                                                                      240
gaaaacccat gtttgtaatt ttaataacat ttttcccatc ttgtaatatc cagagctact
                                                                      300
ttataaattc tctgaaccaa aagtattttc ctcagtgtat ctcttctccc ccagccccta
                                                                      360
                                                                      420
ttgggaaaaa ttacccagta tagttcaggt tatgaggagg atcagccaca caatccagtg
                                                                      480
cttcagtttg aaaatgtaaa attctaaccc taaagtaggg ttggttgaaa tttcagacaa
agcaaaccca gcaggtataa aaagtagtat aaatacaaat ctgtaagtta tttttgaatt
                                                                      540
ttctgaactt ttttctaaga gattacatag gagactaaag aaatctatct gttcaagttc
                                                                      600
taattaggat gattgttaat actgcactgt ggatgaagtg gcgactggct tgtgtgctga
                                                                      660
cttctgtggt ttagcaagag gtttattgtt atcaaatgct aattggcaat gccaagtcac
                                                                      720
tgggaccaat tttctgtttt ataatatcta agtttagaac agaatatata cctgaactgt
                                                                      780
agtggtttga tcggatggag acagaaaacc cgatttttat tctcataaat tttgtggtta
                                                                      840
tttatacaag ggctgtgcta tgctaccata ttcttgttca ataataatag gtttgttgtt
                                                                      900
                                                                      960
ttttttacat tgttaaatgt tccttacccc taaaggtcaa tgttaagtac aacattctga
aaatacaatt tggctacgaa gagtattcat cttctttgaa gctcagtggt tgatatttgt
                                                                     1020
                                                                     1080
gctaataatg caatttcctg attcctgtta caagttatag ctacatatgg gagagactca
gtgagccagc aaaggccata gaaacaacaa tttattaaat gtatttatgg cagaaggacc
                                                                     1140
taaataaact gtgagccacc ttttcttctt tatattgtta catttaagtg ttcttgcttt
                                                                     1200
                                                                     1260
cagcaactca cattaatgct tggagcttat ctctttctct ctctctctc ctctctctct
1320
                                                                     1380
atatatccac accaacatgg gtgacgataa ttcaaagtca tattttgcct ctaagcttga
tcatgttacc tttatgatta aagtatcatg ttatttagcc aatgcaaatc tgttttaaaa
                                                                      1440
caaatagttt aaaaaaagaa caagttttta agggctttat tatagaagaa gtattaatga
                                                                      1500
aggactttcc ttcctccctc cctttcctcc cctccctgcc tcccttcttc ccttccatct
                                                                      1560
cccctcctc cctgccttct ttgtttctcc ttcccttatt cctccctccc tcctttctcc
                                                                      1620
                                                                      1680
cttccttcct ttcttccatt catccttcct tgccttttat ttttatttt tgtaatatca
                                                                      1740
catgtgctgt agtttggaat tttattctag tgcatttctt gctcatcaga acctcagcta
                                                                      1800
atctacctag gaaaaatagt atcaaaggaa atgagaaagt tgtatctgag tccctccaga
                                                                      1860
actaagataa ttctttttga ccatttaagc ctttataaat gcgttttgac catttaagcc
                                                                      1920
tttataaatg cttgttttag gaaagtgaat ctgttagatg catcaacaaa taatgaccag
                                                                      1980
gacaaaacga tttaataatt aaagtctcaa atcaccatgg ttatacattt tcaccagaaa
tagtaatctt acaatttttc atttttctga tgaagatttc tgttccaata tctgtttcct
                                                                      2040
                                                                      2100
aatagatttt ttaaattaat tagctttcct ctgctttatg accacaggtt ttatccctaa
ccgagacagc tgtcttatat ctgcatgcct tagactgtgt ggagggactc catgaagaaa
                                                                      2160
gaccataggt tagaaaaata actcatagta tataccctag taagtgggtt agtagaatct
                                                                      2220
                                                                      2280
cataacatgt attaaaaaga ggttttcttc tctgcttgtt tgtgtcacta gagcaaaatt
gtagagataa tgctcataat gcagtaaata tcagaataat ctacaatatc atttgtggat
                                                                      2340
                                                                      2400
ggtcccaggt cccagtgctc tagttacttt acttcttttt ttttttttga gatggagtct
                                                                      2460
tgctctgtct ctcaggctag agcagtgtgc gatctcagct cactgcagcc tccacctccc
                                                                      2520
aggttcaagc gattctcctg cctcagcctc ccaagtagcc aggattacag gcaccctcca
                                                                      2580
ctaggcccgg ctaattttt ttgtattttt ttagtagaga tggggttttg ccatgttggc
                                                                      2640
caggctggtt tcgaactcct aacctccagt gatccacctg cctcggcgtc ccaaagtgct
```

			•			
aggattacag	gcatgagcca	ccacatccgg	cctaattact	tctttaatcc	ccatttattt	2700
ttatgccatt	ctagcctcat	ttattaataa	aattatgttt	ttactttctc	tttcaggaaa	2760
ttttttaaat	taatattta	tatctagatc	taatgctatg	gaaaagtgcc	tttttatcat	2820
ttataatttc	atttttcact	atttccaaaa	acacataaac	aaatagtttc	agtaggtccc	2880
agcttttact	ttttccattt	aaaccttctt	ttctccattt	cttccctttg	gcttaagaat	2940
aaaagaaaag	gtacattgct	agaattgttt	ctttgggaga	gggtaaaaga	ttacagaatt	3000
agactgttca	gcctttatat	aaactaaatt	tgtcttcatc	tcaaccagct	aatggtaggt	3060
cttatctgaa	tactcatgag	aattttagca	tctgtgaaac	tccatgcacc	agatgtgtgt	3120
aaatttcagg	aagaaagtgt	tgaaagcatt	ttctctgatg	ttaattagat	ggaaataaat	3180
cactaaaaca	tagtttaggt	aaagcctgat	tatgccactt	ttttttaact	agacagggca	3240
aagttgttta	tgttagtgta	cttcttgtct	atcctcagtt	aatttaccta	gacaaaaagt	3300
gtcaaaggaa	atgagaaaaa	ggttatatct	gactccctcc	agacctaaga	taattccttt	3360
tgatcagata	cagtcagatg	gagtgccttg	gtttttgtta	attttgcctc	tattccagct	3420
ccttaccaca	gcggtggtgc	ttaaagaaag	gatcatcagc	aacaggtcag	gatagttcta	3480
cctttgggat	agggctgctt	tccccgtgct	agtatttctg	tgactgttag	tggcactgag	3540
gactgcaaac	ttttatgcaa	tattcttaat	accctattga	tattatgcac	rtcaaccacc	3600
ccaaagaagc	caagaatgct	gtatagtgat	gattccttcc	taatgaatte	atettaaeta	3660 3720
tttagaatgt	tatgtccctt	ttcttttgga	tagccaactt	ggtataaatg	rtacacggac	3780
ttttctaaaa	tgactatata	ggacttaaga	ctttgaaatg	taatttactt	ataaggggaa	3840
ataattatgc	tttagcacat	cattttagaa	acgtcacatt	ttagaaacat	ceagettyet	3900
aacctacatg	tttgggaatt	cattaaaacc	agetgeetat	tananagas	gattataaa	3960
		ttttctacat				4020
gtatggttct	cagcccacca	gctagtatca	graceacery	tanastaga	ccccccaat	4080
attetttgge	cccatcccag	acatactgag	ccayaaactc	tapactttac	gaaccactgg	4140
ctgttttcac	aagccccca	ggtgattctg tagagttttt	atguatatte	cacagettag	gratatacatt	4200
getaagaete	tanattana	ttaacctggt	tatctatea	cttttattgg	caaaagttaa	4260
ttctcactac	tacatataca	gatacagtgt	atttatata	catacacaat	tagtctaatt	4320
cttcataatt	cacttaattt	agtttggcat	tttcctacca	cttactaaaa	ggtttacatt	4380
aaatgactga	tttaaatata	taggtgcaat	gttctatgtt	tattttaatt	gttatgacat	4440
ttaagtaget	aatataatto	accggtgcta	aagteteetg	tttatccata	aaatgggtac	4500
attatagaca	atataataca	agctttcttt	tcattoccta	gtactttacc	agcagaccac	4560
agttttgccc	tagctagacc	aaccctcaga	acaaaatcat	cattccttgt	atttatattt	4620
gtatctgaga	tagtaaacaa	gatggctggc	caggtcaaca	tggcacctta	acttatttt	4680
ttaataggta	aaacttcttc	aaaagtagct	tgctttgtat	aagaactaag	ctatcagtat	4740
agatataget	atccttggag	cttatgtttc	agacaagaat	tatttactaa	aataaataat	4800
aaacaagata	atgcattata	caatttgggc	atttctcgtt	tctcaagtgt	atgcatcatg	4860
gtaaatataa	actaaccaca	agataggtag	attgattcat	ttcattttaa	tctccttgtg	4920
taattcagta	cctccataat	tgttctaatc	ttcttcccac	tgtttacaaa	ttaccagtta	4980
attaactcgt	gaaagaaaaa	ttcacatatc	agaataaaaa	taaatgtata	ctcactttat	5040
		ttccttaata				5100
acaaattttg	gtgctggcaa	atacataggc	aaactgttgg	gagctgctct	agttacattc	5160
ctcccttctt	attccctttt	tctcttcctc	actttattgc	ataacatatt	cctgtaccca	5220
aagcattcta	ccacagttct	atttgactcc	cacttgtaat	aactccttta	aaaaattcca	5280
tgtttaacca	tatgaccctg	cttgcttact	catattctcc	ctccctctcc	ccttcctttc	5340
tctctcttcc	agaagtcatt	tgcctggttt	gaaatatttt	gtagggattg	cttattatat	5400
tattttagct	gatgaacctc	aggacaacgt	ctacacacac	acacatacat	acacgcacac	5460
aaaatctcag	ctgttgaaga	gtgggcttgg	aatcagactt	ctgtgtccag	taaaaaactc	5520
ctgcactgaa	gtcattgtga	cttgagtagt	tacagactga	ttccagtgaa	cttgatctaa	5580
tttcttttga	tctaatgaat	gtgtctgctt	accttgtttc	cttttaattg	ataagctcca	5640
agtagttgct	aattttttga	caactttaaa	tgagtttcat	tcacttcttt	tacttaatgt	5700
tttaagtata	gtaccaataa	tttcattaac	ctgttctcaa	gtggtttanc	taccattctg	5760
ccatttttaa	tttttattta	attttatttg	cttgagcaca	ctgatcaacc	actgaactgc	5820
cttcttccat	tgtcctgcaa	tgatataagg	gttacatttt	tgtgtatatg	gctttcatag	5880
ttgggatttc	anagcnctga	taccanatat	tttcagtttg	ttctctgggg	gaacttcatt	5940
tgcatctatg	tttttancta	tctgtnataa	cttgttaaat	attaaaaaga	tattttgctt	6000
		tcgcaactat			ccadaaataa	6060 6107
aacactgaaa	gttttcaaaa	aaaaaaaaa	aaaaaaaaa	aaadaad		9107

PCT/US99/15849

```
<210> 48
<211> 703
<212> DNA
<213> Homo sapiens
<400> 48
                                                                    60.
ccacgcgtcc gcaggacatc gttttctaca tggtggctgt gttcctgacc ttcctcatgc
                                                                   120
tcttccgtgg cagggtcacc ctggcatggg ctctgggtta cctgggcttg tatgtgttct
                                                                   180
atgtggtcac tgtgattctc tgcacctgga tctaccaacg gcaacggaga ggatctctgt
tctgccccat gccagttact ccagagatcc tctcagactc cgaggaggac cgggtatctt
                                                                   240
ctaataccaa cagctatgac tacggtgatg agtaccggcc gctgttcttc taccaggaga
                                                                   300
ccacggctca gatcctggtc cgggccctca atcccctgga ttacatgaag tggagaagga
                                                                   360
aatcagcata ctggaaagcc ctcaaggtgt tcaagctgcc tgtggagttc ctgctgctcc
                                                                   420
tcacagtccc cgtcgtggac ccggacaagg atgaccagaa ctggaaacgg cccctcaact
                                                                   480
gtctgcatct ggttatcagc cccctggttg tggtcctgac cctgcagtcg gggacctatg
                                                                   540
                                                                   600
gtgtctatga gataggcggc ctcgttcccg tctgggtcgt ggtggtgatc gcaggcacag
                                                                   660
ccttggcttc agtgaccttt tttgccacat ctgacagcca gccccccagg cttcactggc
                                                                   703
tctttgcttt cctgggcttt ctgaccagcg ccctgtggat caa
<210> 49
<211> 639
<212> DNA
<213> Homo sapiens
<400> 49
ggcacgagca ttcacaggtt acaaatgctg ctgccaactg tcctggccaa atgactctgc
                                                                    60
atcacaaacc tttccttgca tgtggagggg atggatttac tcagtccaac tttgatggct
                                                                   120
gcatcacttc tgccctatgt gttctggaag ctttaaagaa ttatatttag tgcctatatc
                                                                   180
cttattctct acatgtgtat tgggttttta ttttcacaat tttctgttat tgattatttt
                                                                   240
gttttctatt ttgctaagaa aaattactgg aaaattgttc ttcacttatt atcatttttc
                                                                   300
                                                                   360
atgtggagta taaaatcaat tttgtaattt tgatagttac aacccatgct agaatggaaa
ttcctcacac cttgcacctt ccctactttt ctgaattgct atgactactc cttgttggag
                                                                   420
                                                                   480
gaaaagtggt acttaaaaaa taacaaacga ctctctcaaa aaaattacat taaatcacaa
                                                                   540
taacagtttg tatgccaaaa acttgattat ccttatgaaa atttcaattc tgaataaaga
600
                                                                   639
<210> 50
<211> 867
<212> DNA
<213> Homo sapiens
<400> 50
ggcacgagca ggtactgggt gactgcctgg ctgaggaaaa gttaactaga cacttgggga
                                                                    60
aaggagatcc aagggagtaa gaggcaaaat gcctttgcat gcttttcttc ctatctctt
                                                                   120
                                                                   180
ttotttetet cetteteact etetecette ettecttet teetttetet ttettttt
tttctctttt cccccacctc tctgcctgcc tccttccttc cctccctcc cctcccttcc
                                                                   240
300
tectetete etectteet geettette ettegttetg ceaacttgee agaaggagee
                                                                   360
caagaaaaag cacccagatg cttcagtcaa cttcttagaa ttcttctttt ttttatgttc
                                                                   420
                                                                   480
agaaaagatg gaaattcatt tctgctaaag agaaagaaaa aattggaaga cagggtgaag
                                                                   540
gtgaacaggc ccattataag aaagaaacaa aaatctatat tctgtctaca aggaagcgag
                                                                   600
agagagaaag agagagaaga aagaagttcc aggattctaa tgtaccaaag ggatctcctt
tttcttgttt tgttctgaaa atttcaccaa aagagcacag gagaacatct tggctaattc
                                                                   660
```

```
720
attggcgatg atgtaagaaa actgagagaa atgaaagaaa tgaagaatta ctgctgcaga
taatatacag ccttgaggaa agaaaggctt ttaagattat agatataaag gctattgctg
                                                                      780
                                                                      840
tattctggga taaaagaaag tctgatgtca gggaaagggg aagttggaaa aactggaaaa
                                                                      867
agaaaaaaga aaaaaaaaaa aaaaaaa
<210> 51
<211> 1569
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (341)
<223> n equals a,t,g, or c
<400> 51
gtattggcca ggctggtctc aaactcctga cctcgtgatc cacccacctt ggcctcccaa
                                                                       60
                                                                      120
agtgcagaga ttacaggcat gagccactgc acctggcctc aagaaaaatt atatatcacg
                                                                      180
tggaatagga tagtagtctc tgcactgatt ttcgttgata atggctgttc ttcttatcac
                                                                      240
cattttgcta tttctttgtc tgggctatta cagggttatt acagaaattt ccagaaagac
                                                                      300
ccctgcctgt cgaatgttta cttcaagctt gagctcctgg tatattatga ggaaattata
                                                                      360
tgatacccca ggagaggtct tcctttccca tgccattgta naattcctaa agtaaaatta
atttgccttc ttgtcaaaga aggagccaat gttgttttaa aattttagct tgagagatag
                                                                      420
gtggggaaga aattaaatag acaagtaatc mctattcaga agagaaggga gagtcattgt
                                                                      480
acgaggccca agatacttgc ccaaaaatat cgcagagaaa aactagtctt tggggtccta
                                                                      540
ttttttgagt ggaacatttg agttatttaa aattagaatt ttattttggt cagattagaa
                                                                      600
                                                                      660
tttctagggt atgtcatatg tgtttttaaa ttgaaagctc ttaaaactcc tattgtagtt
                                                                      720
taatgtcatt atccattaat ttacataaat ctgatttgga tctctatttt catcgtagac
                                                                      780
tgtgtagggg caatttttcc taaaggttct gtgacatagt gctacctttt ttttaaaacc
                                                                      840
tgtcttgccc aggcattatt gagtgccccc tggtgccagc atgtgtattt cacgactgta
tcaacaaatc atgatcatct tctctggcca ttgtgccctt tcagattcca aacttgttac
                                                                      900
ctctcagtcc ttcctacaaa cttagaaagt ctaatatctt aatgtttact tatgtagcaa
                                                                      960
cctccctttc tcccatccct aaatcctctt gtaattaatt attttccttt ggaacttttt
                                                                     1020
aaatctacaa tttccttata atatggtaac caatattaat tttcttgttc tgcgccaagt
                                                                     1080
ttgattttat acaaattgtt tccagtttgg gtcatgagca caaaaccagg tatttttaaa
                                                                     1140
aatctatata acccttcaat gaggcagtat taattttatt aactcattaa ttcaaccaat
                                                                     1200
                                                                     1260
aattettgat tgtttactgt gttagatatt ggggtateee caatacetga cagetgtgag
caaaacaaat gccctacaca catgaggtgt acagtccagt agaaaagata aacaataagc
                                                                     1320
aaattaatag ataatatgat gtccaataag gacttcaaag gaaaataaag cagagtaaag
                                                                     1380
agccagagaa tgacagtgag ctgtttttca catgagtcat cagaaaaggc ctctttaaag
                                                                     1440
aattgacatt tgaacagaaa aacgaatcaa gggcgtcaac tgtttattgc ttttattgct
                                                                     1500
                                                                     1560
1569
aaactcgag
<210> 52
<211> 1196
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (590)
<223> n equals a,t,g, or c
<400> 52
gattggttct gtttatgtga tagattactt ttattgattt gtatgttgaa ccagccttgc
```

```
atcctaggga tgaagccgac ttggttgtgg tggataagct ttttgatgtg ctgctgggtt
                                                                        120
tggcttgcca gtgttttatt agggattttt gcgtcaatat tcatcaggga tattggcctg
                                                                        180
gaattttctt tttttgttat gtgtctgcca ggttttggta tcagggtgat gctggcctca
                                                                        240
taaaataagt tagggagggc tccctcttt tctttcattt ggaagaattt cagaaggaat
                                                                        300
ggtaccagat ccyctttgta cctctggtag aatttggctg tgaatccatc tggtcckgag
                                                                        360
ctttttttt gttggtaggc tattaattac tgcctcaatt tcagaacttg ttattggtct
                                                                        420
                                                                        480
attcagggat ttgacttctt cctggtttag tcttgggagg ttgtatgtgt gcaggaattt
attcatttct tctagatttt ctcgtttatt tgtgtagagg tgtttatagc atyctctgat
                                                                        540
ggtagtttgt attctgtggg atcagtggtg atctcccctt tatcattttn attgtgtcta
                                                                        600
tttgattctt ctctcttktc ttctttatta ttctygctaa tggtctatgt attttgttaa
                                                                        660
tctyttacaa aaacaggctt ctagattcat ggatgttttg aaaggttytt cgtgtctcta
                                                                        720
totocttcag ttottccctg atottagcta tttcttgtct tctgctagct tttgaaattg
                                                                        780
tttgcttttg cttctctagt tcttttaacc gtgatgtcca gtgtgtcaat ttcagatctt
                                                                        840
                                                                        900
tccagccttc tgatatgggc atttaatgct ataaatttcc ctcttaacac tgctttagct
gtgtcctaga gattctggta cgttgtctct ttgttctcat tggtttcaaa taacttcatt
                                                                        960
atttctgcct taattttgtt atttacccag cagtcattca agagcaggtt gttcaatttc
                                                                       1020
catgtagttg tgtggttttg agtgaatttc ttaatcttga gttctaattt gattgcactg
                                                                       1080
tggtctgaga gacggttaca atttccattc ttttgcattt gctgagaagt gttttacttc
                                                                       1140
caattgtgtc tcgtgccgaa ttcgatatca agcttatcga taccgtcgac ctcgag
                                                                       1196
<210> 53
<211> 945
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (295)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (875)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (914)
<223> n equals a,t,g, or c
<400> 53
gaatggtgaa atattaagtg ctttctcccc caggttcagg attatgacag ctatgtccat
                                                                         60
tcacctcttc tgtacagcat tgtcctgtgg aagttctggc cagtgcaata aggcaattaa
                                                                        120
                                                                        180
aagaaataaa atatcaaacg attggaaaga tgttaatgtg tcatcattca tagaaaacat
                                                                        240
gattcataga tatacataca cgaatgcttt gaattcataa gtagattcag ccagttgctg
                                                                        300
gatataaagt caatatacaa aaactatttt tatagacatg aaacacgcaa tgagnaaaaa
aatttaacca tttttagtag catcaaaaaa cccccatacc taggaatatg aatttgtagt
                                                                        360
actatttggg atatgttgat ggatatttat catttccagt ttgggattat tataaagaaa
                                                                        420
atagccctga acatttgtaa tatatgactt ttggtgaatg tagcattcat ttctgttgat
                                                                        480
tacaaactca ggggtgaaat tgttgagtcc taagggagct atagatgtat tcaacttcag
                                                                        540
ctgatatggc taaataaatt tgcgaaaaag attgcatcaa gttatgctcc catcagcaat
                                                                        600
                                                                        660
atgagagttc ctgtttttcc acattgtcag caacactttg tactgttact ccttttaatt
ttagccgatt tggctgaagg tgtggtaata tctcattgta gtggccaggc gtggtgctca
                                                                        720
cgcctgtaat cccagcactg tgggaagcca aggtgggccg atcacgaggt caggagatcc
                                                                        780
agaccatect ggetaacatg atgaaaceet gttgeetgta gteecaacta ettgggagge
                                                                        840
tgaggcagga gaatggcatg aactcgggag gcggngcttg cagtgagcct ccagcctggg
                                                                        900
                                                                        945
caacagagtg agantctctc aaaaaaaaaa aaaaaaaac tcgag
```

PCT/US99/15849 WO 00/04140

31

```
<210> 54
<211> 488
<212> DNA
<213> Homo sapiens
<400> 54
ggcacgagga gagtagaggc tattcatgta atgtctataa aaaaataaca ccaaggctgg
                                                                        60
gattacaggc atgagccact gcacctggcc agtttgctta ttttgtttgg tgcctcctcc
                                                                        120
                                                                        180
catgggagac ctcaaggagg tatgcctgcc ccacagatgc cctggaagga cagcttgctg
ctcctactca gaaccacacc tgcagacaga ggaggacaga cggacactca tttgctgagc
                                                                        240
                                                                        300
acccatgtaa catgaactaa gagctgggtg gagacaatga acggtggagc catcgttccc
                                                                        360
gatgtggagg gagaacagct caagaccacg gaacagcctg ctctcccgct tcctggcttc
                                                                        420
cgtgcgcttt tgtccaatca ggctttttga ccaatcggcc aggcgcgcta tgtaaatttc
                                                                        480
tgacattttc aaagctgtct ttttaataaa cctttcagtg taaaaataaa aaaaaaaaa
                                                                        488
aaaaaaaa
<210> 55
<211> 2860
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (753)
<223> n equals a,t,g, or c
<400> 55
ggcacacagg gctggcaggc ccgcggtggc tggtgttgag gcatgaacaa attgtaccgg
                                                                         60
                                                                        120
gtatccccca ccccactctg accaccagtt cctccttgga tatcactccc cctgacaggc
ageccaecca ggeetggatt tgteeetgte teeccetttt getttteece catgactaat
                                                                        180
gggcaccagg tettgetget cetgettete acetetgeag tggcagcagg cecetggece
                                                                        240
caggtgcatg ccggtcagtg gggttggatg tgccttcctc caggcctgcc ctctgtccaa
                                                                        300
                                                                        360
gcccggagtg ggcttggtgg gctccctggt ggcccccagt gggtgccagg tggtgcccgg
                                                                        420
ggttattgag gggtggttgt atcactgtag ggacaggctt cttgccccag cctggagagc
                                                                        480
tgttttcttc aggaaggttc tggagatgga gacttgtttg cgaattcacc acaactccag
gctgggaggc tgggtctctg ctctcagagc cgagacacca gggaggatag ccaggctgcc
                                                                        540
                                                                        600
ctgcctggga attctgctgg gccgtcaaat tcaacccgca ccaacgtggg caggaggcca
cagtgtcctg ccaggagcag agggctgaag gtctgcagga ggaagaccct atcctggtgg
                                                                        660
                                                                        720
ggggcacctg ctgcccaccc tgccccagc gtgcctgggg ggagcacacc tgggcatgga
                                                                        780
ggagtccagg gtgctgggcc acacaagaga ggngggggag aggcctggac agtaggaaga
                                                                        840
tottgcccag ggtcctggat ccgccactct gggggtgacc ttggacaaac ctctgccttg
                                                                        900
gccctcagtc tccccatcaa ggtttttcca ttcaggaggg tttgggccat cttcagccac
cctaccagcc ctgaaaagga tgtgactcct gtttctggga agtgtgtggt gtgttaggtg
                                                                        960
                                                                       1020
ggcctacagc cctggttgtg gggagggaag gatggagaga cagcacagtg acagagccca
                                                                       1080
gactgcaggc tggagtgagg gttccacttc cccgctgctg tgtgtcctgg accagtgcct
                                                                       1140
ctgaaccttg gcacttgggg cagtggatat taacatcttt ccaagcccaa ttcttggggc
                                                                       1200
atcagggeet ceggteetet gggaggtgge aggteeteag attggagatg ceatgggggg
                                                                       1260
gggaggtgcc tctcctttgg agggtatgga agtggagaca ggagtggcct ggcgcagctg
                                                                       1320
ccgtggttct taggggctgg gcccggggag cccatggggc ttgtgcctag aaagcctggg
                                                                       1380
ctcctcactg gggtctagat gtgcagactt catgtctccc cagctccagc tctgttctct
ataggtcaag cctccacaat gccagaggcc cagggctagc cccctccacg tccctcctag
                                                                       1440
                                                                       1500
atctacaget geceetttga tgacagegee attgagteee etgggetggg ggggteatge
                                                                       1560
aggggtgagg cagctgcctg ccgccggtac tcattgcctg gccaggcagg acacaggctg
gcgggcactg agagtgggcc ccacgaaatc cattgtcagg ttaccaggat gaagaaccca
                                                                       1620
```

ggctggtcgt ggagtgcagg gcggggcctg ccggaagaat tatgggcact gcagcaggag

PCT/US99/15849

	1740
ggcagcctgg gccattagct cctgatgtca tcgatttggg tgaggggaca gggaagtcag	1800
aggaagetgg ccagtggete teacgeagae ttacageagt ggagtggtge etgatteetg	1860
gtacagetge teceaetgag tetecaggga tetgtggtte aggaceeet geaaeteet	1920
cccagaccc tgtactggtg ggaggagagg acctagagga aaggtgctgg gcagacaagc	1980
agctgaggga ggccctgggt ttagcttatc agtcttctgg gccctcctgc cccaggaagg	2040
gragggagga ccatggtgtt gcccctgtca tcgttatcgt cctggccatg agcttgdayy	
actgggaggg ccggagtcag ccaggcagac ggcagcacag catttgcctg ttggcaggtg	2100
gccttggtgg cttcccaaag gcaatcgctc cacgcagaac aaaactcact tttttggggg	2160
gtgaagcacc ttggttcatt tgtttagttc gttaattcca gcagtctgtt tctaagggaa	2220
acatogotoc agoogotoct gogootocca cootoccaco aggigoccag igitoccaag	2280
ggcccgaat cccaacctta ttcaggcgtc agcatctctg cacccaaat gcctgttagg	2340
gaggatagtg aaggetgage ceteetggge ceateaaaag ceageagtga gagaacaeee	2400
cratchetet gaggigacet igtagggeag teegigeigt eiggeiggee egggegagge	2460
gggcagggac caaggcctgg cgcctgggcc tcgctggcct tgctctgcgt gctgactta	2520
tectgatagt accttgattt tectactgtg actteceett etgtegaett eeleaceae	2580
tttaaaattc cgtattgaga gcagtttcct aagttacctc aaatcctatt cagaagaayy	2640
ttcttcctgg aagttgggag ggcggaaaac aagtttagtc acagaagact actccatgtt	2700
tgagcttctg tttcaaggga agtgagtaac tgccggagga gccctgccc tctgcagtgt	2760
gtggtgttgc cctgatactt ttcagattga ggtgttactt acatgtaata aaatgcacag	2820
acttaagtgt aaaaaaaaaa aaaaaaaaaa aaaaaaaaa	2860
<210> 56	
<211> 1559	
<212> DNA	
<213> Homo sapiens	
various daptone	
<220>	
<221> SITE	
<222> (1445)	
(222) (1445)	
2223 n emple art d or c	
<223> n equals a,t,g, or c	
<220>	
<220> <221> SITE	
<220> <221> SITE <222> (1551)	
<220> <221> SITE	
<220> <221> SITE <222> (1551) <223> n equals a,t,g, or c	
<220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56	60
<220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cqtqctgggc aggcgctgcc gagacettct cctgcagetc	60 120
<220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagacettet cetgcagete tacctacage ggcggaget gegggtgcc gtgcetgagg tectactgca cagegaaggg	
<220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagacettet ectgcagete tacctacage ggccggaget gegggtgece gtgcetgagg tectactgca cagegaaggg gctgccagea gcagcgtctg caagetggae ggaeteatee accgetteat cacgeteett	120 180 240
<220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagaccttct cctgcagctc tacctacagc ggccggagct gcgggtgccc gtgcctgagg tcctactgca cagcgaaggg gctgccagca gcagcgtctg caagctggac ggactcatcc accgcttcat cacgctcctt gcggacacca gcgactcccg ggcgttggag aaccgagggg cggatgccag catggcctgc	120 180 240 300
<220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagaccttct cctgcagctc tacctacagc ggccggagct gcgggtgccc gtgcctgagg tcctactgca cagcgaaggg gctgccagca gcagcgtctg caagctggac ggactcatcc accgcttcat cacgctcctt gcggacacca gcgactcccg ggcgttggag aaccgagggg cggatgccag catggcctgc ggaagctgg cggatgcag catggctgc cggaagctgg cggatgcag catggctgc ctgctcaggc acctgccat gatcgcggcg	120 180 240
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagacettet cetgcagete tacetacage ggccggaget gegggtgece gtgcctgagg tectactgca cagegaaggg gctgccagca gcagcgtctg caagetggac ggactcatec accgetteat cacgetectt gcggacacca gcgactcccg ggcgttggag aaccgagggg cggatgccag catggcctge cggaagetgg cggtggcga cccgctgctg ctgctcagge acctgccat gatcgcggcg ctcctgcagg gcggaccca cctcaacttc caggagttcc ggcagcagaa ccacctgagc</pre>	120 180 240 300
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagaccttct cctgcagctc tacctacage ggccggagct gcgggtgccc gtgcctgagg tcctactgca cagcgaaggg gctgccagca gcagcgtctg caagctggac ggactcatcc accgcttcat cacgctcctt gcggacacca gcgactccg ggcgttggag acccgagggg cggatgccag catggcctgc cggaagctgg cggtggcga cccgctgctg ctgctcaggc acctgccat gatcgcggc ctcctgcacg gccgcacca cctcaacttc caggagttcc ggcagcagaa ccacctgagc tgcttcctgc acgtgctgg cctgctggag ctgctgcagc cgcacgtgtt ccgcagcgag</pre>	120 180 240 300 360
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagaccttct cctgcagctc tacctacagc ggccggagct gcgggtgccc gtgcctgagg tcctactgca cagcgaaggg gctgccagca gcagcgtctg caagctggac ggactcatcc accgcttcat gcggacacca gcgactcccg ggcgttggag aaccgagggg cggatgccag catggcctgc cggaagctgg cggtggcga cccgctgctg ctgctcaggc cggatgccag catggcctgc ctcctgcacg gcgcaccca cctcaacttc caggagttcc ggcagcagaa ccacctgagc tgcttcctgc acgtgctggg cctgctggag ctgctgcagc cgcacgtgtt ccgcagcgag caccacgtgggg cgcaggagg cctgctggag ctgcttcatcc gcctgctgct gaattacagg</pre>	120 180 240 300 360 420
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgtgcc gtgcctgagg tcctactgca cagcgaaggg gctgccagca gcgggtgcc gagactcatc accgctcatc accgctcaga gcggacacca gcgactccg ggcgttggag accgagggg cggatgccag cagcgtcgc ggcgttggag accgagggg cggatgccag catggcctgc caggagggg cggaagctgg cggatgccag catggcctgc caggagggg cggaagctgg cggatgccag catggcggc ctcctgcacg gccgcacca cctcaacttc caggagttcc ggcagcagaa ccacctgagc caccaggggg cgctgtggga ctgcttctt tccttcatcc gcctgctgt tccttcatc accagtgtt tcctacaggagg cgctgtggga ctgccttctt tccttcatcc gcctgctgct gaattacagg aagtcctccc gccatctggc tgccttcatc accaggttt tcctacagactcaagactaagactcaagactaagactcaagactcaagactcaagactcaagactcaagactcaagactcaagactcaagactca</pre>	120 180 240 300 360 420 480
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagaccttct cactacage ggccggagct gegggtgcce gtgcctgagg tcctactacage gctgccagca gcagctgtgag aaccgagggg accggaccaca gcagctccag ggcgttgag aaccgagggg cggatcatca cacgctcctt gcggacacca gcgatccag ggcgttgag aaccgagggg cggatgccag catggctgc catgatcagaggg cgaagctgg cccgaagctg cccgaactc catgagtcag acctgccat acctgccat gatcgcggc ctcctgcag gccgcacca cctcaactc caggagttcc ggcagcagaa ccacctgagc tgcttcctg acgtgtgga ctgcttctg tccttcatca gcctgctgt gaattacagg aagtcctcc gccatctgc tgcttcatc aacaagtttg tgcagttcat ccataagtac attacctaca atgcccaag agccatctcc ttcctgcaga agcacgcga cccgctccac acctgcagc acctgctact cataagtac attacctaca atgcccaag agccatctcc ttcctgcaga agcacgcga cccgctccac</pre>	120 180 240 300 360 420 480 540
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gagaccttct cctgcagctc tacctacagc ggccggagct gcgggtgccc gtgcctgagg gctgccagca gcagcgtctg caagctggac ggactcatcc accgctcctt gcggacacca gcgactcccg ggcgttggag aaccgagggg accggatgcag cggatgcag cggatgcag ccggatgcag accggatgcag catggcctgc cggaagctgg cggtggcga cccgctgctg ctgctcagg acctgccat gatcgcggcg ctcctgcacg gccgcacca cctcaacttc caggagttcc ggcagcagaa ccacctgagc tgcttcctgc acgtgctggg cctgctggag ctgctgcag cgcacgtgtt cgcaccagggg cgctgtggag ctgcttctt tcctcatcc gccacgtgtt caccaggggg cgctgtggag ctgcttctt tccttcatcc gccacgtgtt caccaggagg cgctgtggag ctgcttctt tccttcatcc gcctgctgct gaattacagg aagtcctccc gccatctgc tgccttcatc aacaagtttg tgcagttcat cataagtac attacctaca atgcccagc agccatctcc ttcctgcaga agcacgccga cccgctccac gacctgtcct tcgacaacag tgacctggtg atgctgaaat ccctccttgc agggctcagc</pre>	120 180 240 300 360 420 480 540
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc aggcgctgcc gtgcctgagg tcctactgagggggagacag gctgcagagggggagacag gggggtgccc gtgcctgagg tcctactgaa cagcgaaggg gctgcagaa gcagctctc gggacacca gcgactcccg ggcgttggag aaccgagggg cggatgccag catgcctgc cggaagctgg cggtggcga cccgctgctg ctgctcagga acctgccat gatgcctgc ctgctagag ctgcttcatc acgctcctt caggagttcc ggaagctgg cggtgggg cctgctgag ctgctagag ccacctgagc caccagggg cctgctgag ctgcttcatc aggagttcc ggaagcagaa ccacctgagc caccagggg cgctgtggag ctgcttctg tccttcatcc gccacgtgtt ccgcagcgag caccagggg cgctgtggag ctgcttctt tccttcatcc gcctgctgct gaattacagg caccagggg cgctgtggag agccttcatc aacaagtttg tgcagtcatc ccataagtac attacctaca atgcccaag agccatctcc ttcctgcaga agcacgcga ggacggagagcctggacg aggacgagag ggaggagagc cctgctgcaga aggagggagagc ggaggagagagc ggaggagagaga</pre>	120 180 240 300 360 420 480 540 600 660
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc gtgcctgagg gctgccagagggggggggg</pre>	120 180 240 300 360 420 480 540 600 660 720
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag gegggtgec gtgcctgagg tcctactgca cagcgaaggg gctgccaga gcagcgtctg gagactatc ggggagacaca gcggatccag ggggttgaca ggactcatc gggagacaca gcgactccag ggcgttgag aaccgagggg cagtgccag caggagctgc ggggttgaca accgctcat cacgctctt gcggaagctg ggggttgaca aaccgagggg acctgacag aacctgagac ggatgccag catggcaga cactgccag catggagagctg cggaagctg ccccaacttc caggagttcc ggaagctga acctgccag acctgccag gcgaagctg ccccaacttc caggagttcc ggaagcaga ccacctgaga ccccaactgaga ccccaaggag cgctgtgag ctgctagag ctgctgaga ctgctgaga ctgctgaga ctgctgaga ctgctgaga ctgctgaga ctgctgaga cgcaagagaga cacactgagag cacaaggggg cgctgtgga ctgcttctt tcctcatc gacagtgt ccgcaagcaga aagacctcc gacagtgtt tccttcatc gcctgctgaaatcaattacaaagaattacaaattacaaattacaaattacaaattacaaattacaaattacaaattacaaatacaatacaattacaaatacaattacaaatacaattacaaatacaattacaaatt</pre>	120 180 240 300 360 420 480 540 600 660 720 780
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc gggctgcc gctgccagcag ggccgagct gcgggtgcc gtgcctgagg tcctactgcagcaggggagacag gcggctgcc ggggtgccag ggactcatc gcggaagcag gcggttggag aaccgagggg cggatgccag catggctgc gggatgccag ggggtgccag ggcgttggag aaccgagggg cggaagcaga cccgctgctg ctgctcagga acctgccat gatggcggagctgcc ggaagctgg cggaagctgg cgcgaacca cctcaactc caggagttcc ggcagcagaa ccactgagc ctcctgcagg cctcatcaggaggg cgctgagga cctgctgagg ctgctgaga ctgctgaga ccactgagagagagagagagagagagagagagagagagag</pre>	120 180 240 300 360 420 480 540 600 660 720 780 840
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag cgtgctgggc atccagcagt ggggagacag gegggtgcc gtgcctgagg tcctactgca cagcgaaggg gggagacaca gcgggtggc ggactcatc accgcttcat cagcgaaggg gggagacaca gggcttggag aaccgagggg cggaagcag cggaagctgg cggaagctgg cggaagctgg cggaagctgg cggaagctgg cggaagctgg ccggaagctgg ccgaacca acctgaagg caccagaggg caccagaggg ccgcaacca cctcaactc caggagttcc ggcaagcaga ccacctgagc ccgcaagggg caccagaggg ccgcaagag ccgcaagag ccgcaagagagag</pre>	120 180 240 300 360 420 480 540 600 660 720 780 840 900
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag ggggtgcc ggggtgcc gtgctgagg tcactactacagc ggccgagct gcgggtgcc ggggtgcc gtgctgagg tcactactacagc gcggttgcgagacaca gcgactcccg ggcgttggag aaccgagggg cgggagacaca gcgactcccg ggcgttggag aaccgagggg cggatgccag caggtgcgca ccgctgctgc ctactgcaag gcggtggca ccgctgtgag aaccgagggg cggatgccag ctgctcatacacaggggg cgcgaagctg ctgctgagg ctgctgagg ctgctgagg ccgctgagg ctgctgagg ctgctgagg ccgctgagg ctgctgagg ctgctgagg ctgctgaga ccactgagg cgctgtgag ctgctgaga ccactgagg cgaagagaga caccaggggg cgctgtggag ctgctgaga ctgctgagag ctgctgaga ccacctgagg caccaggggg cgctgtggag ctgctgagag ctgctgagag ctgctgagag ctgctgagag ctgctgagag ctgcttcatacaagg ctgctgagag ctgcttcatacaagg ctgcttcatacaagg ctgctgagagagagagagagagagagagagagagagagag</pre>	120 180 240 300 360 420 480 540 600 720 780 840 900 960
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag ggggtgccc gtgcctgagg tcctactcacagc ggcggagct ggggtgccc gtgcctacacacacacacacacacacacacacacacacac</pre>	120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag gggggtgcc ggggctgcc ggggctgcc gggggggg</pre>	120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140
<pre><220> <221> SITE <222> (1551) <223> n equals a,t,g, or c <400> 56 atccagcagt ggggagacag ggggtgccc gtgcctgagg tcctactcacagc ggcggagct ggggtgccc gtgcctacacacacacacacacacacacacacacacacac</pre>	120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080

```
1260
                                                                     1320
ccggggatcc tcgaggcaaa gcccaggaag cgtgggcgtt gctggtctgt ccgaggaggt
                                                                     1380
gagggcgccg agccctgagg ccaggcaggc ccaggagcaa tactccgagc cctggggtgg
                                                                     1440
ctccgggccg gccgctggca tcaggggccg tccagcaagc cctcattcac cttctgggcc
acagneetge geggagegge ggateecee gggeatggee tgggetggtt ttgaatgaaa
                                                                     1500
cgacctgaac tgtcaaaaaa aaaaaaaaaa aaacccgrgg gggggcccgg nacccaatt
                                                                     1559
<210> 57
<211> 2064
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (2001)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (2024)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (2049)
<223> n equals a,t,g, or c
<400> 57
                                                                       60
atgggcgagg ctgcggggcc ccggcgcgca cgcccgcacc tctccccagc cctggcgtgg
gcccagcccg gcccaggcag caatggggtt cctgcagctg ctggtcgtar cggtgctggy
                                                                      120
atccgaacac cgggtggctg gtgcagccga ggtcttcggg aattccagcg arggtcttat
                                                                      180
                                                                      240
tgaattttct gtggggaaat ttagatactt cgagctcaat aggccctttc cagaggaagc
tattttgcat gatatttcaa gcaatgtgac ttttcttatt ttccaaatac actcacagta
                                                                      300
                                                                      360
tcagaataca actgtttcct tttctccgag gcgtagatcc cccaccatgt gacgctggga
                                                                      420
cagaccagga ctccaggtgg aggttgcagt atgatgtcta tcagtatttt ctgcctgaga
atgacctcac tgaggagatg ttgctgaagc atctgcagag gatggtcagt gtgccccagg
                                                                      480
                                                                      540
tgaaggccag tgctctcaag gtggttaccc taacagctaa tgataagaca agtgtttcct
tctcctccct cccgggacaa ggtgtcatat acaatgtcat tgtttgggac ccgtttctaa
                                                                      600
                                                                      660
atacatctgc tgcctacatt cctgctcaca catacgcttg cagctttgag gcaggagagg
gtagttgtgc ttccctagga agagtgtctt ccaaagtgtt cttcactctt tttgccctgc
                                                                      720
                                                                      780
ttggtttctt catttgtttc tttggacaca gattctggaa aacagaatta ttcttcatag
                                                                      840
gctttatcat catgggattc ttcttttata tactgattac aagactgaca cctatcaagt
atgatgtgaa totgattotg acagotgtca otggaagogt oggtggaatg ttottggtag
                                                                      900
ctgtgtggtg gcgatttgga atcctctcga tctgcatgct ctgtgttgga ctagtgctgg
                                                                      960
ggttcctcat ctcgtcagtg actttcttta ctccactggg aaacctaaag atttttcatg
                                                                     1020
atgatggtgt attctgggtc actttctctt gcatagctat cctcattcca gtagttttca
                                                                     1080
tgggctgcct aagaatactg aacatactga cttgtggagt cattggctcc tattcggtgg
                                                                     1140
                                                                     1200
ttttagccat tgacagttac tggtccacaa gcctttccta catcactttg aacgtactca
                                                                     1260
agagageget caacaaggat ttccacagag etttcacaaa tgtgeetttt caaactaatg
                                                                     1320
acttcattat cctggcagta tggggcatgc tggctgtaag tggaattacg ttacagattc
                                                                     1380
gaagagagag aggacgaccg ttcttccctc cccacccata caagttatgg aagcaagaga
gagagcgccg agtgacaaac attctggacc ctagctacca cattcctcca ttgagagaga
                                                                     1440
ggctctatgg ccgattaacc cagattaaag ggctcttcca gaaggagcag ccagctggag
                                                                     1500
agagaacgcc tttgcttctg tagatgccca ggggcttggt cagtgtgcct cagctttgga
                                                                     1560
                                                                     1620
gttcatgcct ggagtggttc aacagtctct ggtgcaagtc taataagaga tcaggcatat
                                                                     1680
atatctgttc tttgcataat attatggtgc ccttattgat atatggtaag ggtgtactag
                                                                     1740
gggattagga tgattgtaag agaatgagaa agatgaccaa aaggttggtg gtagggaggc
```

			34			
		ttgagaaatt ttaaaaaaat	taaaaactya	CCCCCCGCC	5-550	1800 1860 1920
gtttctttt	ataacccctt	gaaacaagtc	Coccacciga	estt sagga	rrtgatactt	1980
	~++~~~~+c++	tatcrgtgac	adtalligga	gacccaggg	• • • •	2040
aggcctttga	attttagaat	ncaaaaaggg	gagccagcca	gacngggggg	Ccaaccegga	2064
tccccaacng	ggaccagggg	ggtc				2000
<210> 58 <211> 1050						• .
<212> DNA						
<213> Homo	sapiens					
(223)						
<400> 58					carcatoro	60
ccacgcgtcc	ggccagccag	teegeeegte	cggagcccgg		cagcatggcg gctgctgctg	120
	. +-a+a+aaaa	- accacaaacc	aaaaacgccg	gccccccgg	95-5	180
		CCCCCCCCCCC	Lacacacaca	Cgg caaagg-	5	240
	~+ ~+ ~ ~ ~ ~ ~ ~ ~	cttaactaaa	CEGGCGCCCC	ccgccgccc	099-33	300
		addacaatac	adaccuqueg	cgggcgccg	- 3 -	360
	. ~+~~~~~~~	, aacacaaaca	gaddcdcaga	ggccgaggac		420
	. ~~~~~+~~+	, cacatctaaa	acaccccccc	Caaccccgac	0099	480
		· acacctacac	COCACCCCC	, tegegeeee		
		, acadeceade	. Efatececa	, gcccgcccc	95-55	540
		· tacdacdacd	raccccucuu	CCCGGGGCGG	9-33-33	600
		r daccccdado	: Lulluayyu			660
		· aaaataacac	r ccccacacac	1 666666	J J -	720
	- atatasacto	r cccctgagg	i acatacian	, ggcgccgcc	· -9-9-5-	780
		- caddtdcct	T Cacucucucu	, cccgcaaa,		840
•		, <i>«»сссапаа</i> (i taccccuci	, attecegees		900
ccggatcccg	g tgcaccetg	g gacccagaa;	caaccaacc	gccctctcac	ccgaggatcc a aaaaaaaaaa	960
ccccgccago	acglecaya	t assestant	r tmaamcamca	a aaaaaaaaa	a aaaaaaaaaa	1020
ctaccccctg	g gccccacaa	L aaacacgac	o tgaagues.			1050
aaaaaaaaa	a aaaaaaaaa	a aaaaaaaaa	4			
						•
<210> 59						•
<211> 253	3					
<212> DNA						
<213> Hom						
	-					
<400> 59				teatat	t taccttcatt	60
ccacgcgtc	c gcctggcaa	c ccctaatat	t tgggatete	t caatgetat	t tgccttcatt	120
	~ ++ -+ ~ - + + -	c cacttdgtg	a attucutu	c cccggccgg		180
_ 4_ 4 4 4_ 4_ 4_		t cataadadc	t ttqaqatta	c ggaggacas		240
		a cattcattt	a daadatatu	g ccacaaaca	9 009-9	300
	~ +~~~~~	c tttacqtct	c attcaaqaa	a ccyaagegu	C 0000-5-55	360
	tasataata	r ffaccatt	t datadaycu	g gacagears		420
	- ++ccccsssc	re tatetacco	a acticidayo	ig ccaaces		480
	_ + -+ -+ -+ ~ -	t daaaaacta	C CCCCCGaa	ic cegagagag	u	540
		o ttttaaag	ia ctadaccii	.q gacccages		600
		a ctttacada	it ducticaqu	,c cgcccgcac	· · · · · · · · · · · · · · · · · · ·	660
	aaataacac	a dadttcada	id ffcffcage	ic ggttageet		720
	·+ asactecte	an accettact	t actccaqca	ic cccacce	.c. cog	
	atassaata	-t acctcator	t cattetge	ic geeeggaag	u 9000	780
	rotatoggt	e ctttgaaag	t cagcaccag	ic cagcaces	u gegeen	840
	a agaagtea	ar araactdaa	at aatqttcaq	a caycagege	.g cage-s-	900
		-t daatdaggi	ta ataatici	ty aayatguu.	,	960
		amaactadi	ra cadaduc	LL accessor	·	1020
geeegeact	ra ttcacccc	ca cottcaao	ca agcaacaa	tt gctgggaag	ga ggccatttct	1080
ccaaayct	ga ccoageoo.		-			

			33			
caggtcgaca a	actoctaco	aagaaataca	gataaaaaag	gcaagcctga	aatagcatgt	1140
gaaaacccac a	ttgtacagt	aagtaccttt	gaagcagcct	actctacaca	ttgcagacaa	1200
agatccaatc c	cagaggagc	aggaattaga	agcttatgta	gatgatatag	atattgatag	1260
tgatttcaga a	aggatgatt	tttattactt	gtctcaagaa	gacaaagaga	gacagaagcg	1320
tgagcatgaa g	aatccaaga	gggtgctcca	agaattaaaa	tctgtgctgg	gatttaaagc	1380
ttcagaggca g	aaaggcaga	agtggaagca	acttctattt	agtgatcatg	tgtttcttca	1440
tatagcttta a	aattatgct	attgacatta	tgggaaagat	ttatcaatga	gagaaatgtg	1500
tctctttttc a	gccgtgttg	aaatccttgt	ctcctgtaga	cccagtggaa	cccataagta	1560
attcagaacc a	tcaatgaat	tcagatatgg	gaaaagtcag	taaaaatgat	actgaagagg	1620
aaagtaataa a	tccgccaca	acagacaatg	aaataagtag	gactgagtat	ttatgtgaaa	1680
actctctaga a	ggtaaaaat	aaagataatt	cttcaaatga	agtcttcccc	caaggagcag	1740
aagaaagaat g	tgttaccaa	tgtgagagtg	aagatgaacc	acaagcagat	ggaagtggtc	1800
tgaccactgc c	cctccaact	cccagggact	cattacagcc	ctccattaag	cagaggctgg	1860
cacggctaca g	rctgtcacca	gattttacct	tcactgctgg	ccttgctgca	gaagtggctg	1920
ctagatctct c	tcctttacc	accatgcagg	aacagacttt	tggtgatgag	gaggaagaac	1980
aaataataga a	igaaaataaa	aatgagatag	aagaaaagta	agaaccaaga	ttcatatgaa	2040
gtgatattag a	ttgttcctt	ttacaaaagt	gtttagcttc	aagactggaa	agggaatatg	2100
agtgtaagtt t	actatatat	aaagctaaga	tgtggattta	caggaagaac	cctggtttga	2160
ataactgatc t	gaaattagt	agttacctgt	aaatggcaga	tcttttagga	aaataagaga	2220
aaggtaaggg c	tcttttgaa	taaactgctg	ttttatttgt	ggcacaactg	atcaatcttg	2280
gaaattcttt a	agtatttt	aataagaaat	gaattatcat	ttcttgccag	aatttgctac	2340
cttaaggtga t	tgggaaaat	tctgttgcaa	gaacattaac	atttagtatg	actcctttt	2400 2460
actgtattct t	gcagttaat	aactgcagct	attatgttaa	taacaagttg	tttgtattt	2400 2520
atttttgttt a	ataccagtct	taaagatcca	ggttctgaat	aaaaaacta	attgatacaa	2520
aaaaaaaaa a	aaa					2333
<210> 60 <211> 899 <212> DNA <213> Homo s	sapiens					
<400> 60						60
ggcagatttc c	ccggcacctt	cgtgggcacc	acagagcccg	cctcccacc	cetgageage	60 120
acctcaccca c	ccactgctgc	ggccactatg	cctgtggtgc	cctctgtggc	cagectggee	180
cctccggggg a	aggcctcgct	ctgcctggaa	gaggtggccc	cccctgccag	egggaecege	240
aaagctcggg t	tgctctatga	ctacgaggca	gccgacagca	gtgagetgge	cetgetgget	300
gatgagetea t	cactgtcta	cagcctgcct	ggcatggacc	ctgactggct	cattygegag	360
agaggcaaca a	agaagggcaa	ggtccctgtc	acctacttgg			300
gccccatcc c				tagacacact	accacttaac	420
	ccccgcatt	ctggcctagg	caggagagga	tgggcgcact	gccacttaac	420 480
ttgtttgttg g	gtgacacagt	tgttcagagt	caggagagga ggggagaatt	tgggcgcact caccccattc	gccacttaac tgtccctgcc	480
cctagtcacc t	gtgacacagt tagctgtgag	tgttcagagt ggtgcctgag	caggagagga ggggagaatt gctgaatggc	tgggcgcact caccccattc tccaccctcc	gccacttaac tgtccctgcc cccagccctg	480 540
cctagtcacc t	gtgacacagt tagctgtgag gtggctctgg	tgttcagagt ggtgcctgag agcccctgcc	caggagagga ggggagaatt gctgaatggc cctgcctgca	tgggcgcact caccccattc tccaccctcc tccccgagca	gccacttaac tgtccctgcc cccagccctg ccccaccctc	480 540 600
cctagtcacc t cttctgacct c caggctccac t	gtgacacagt tagctgtgag gtggctctgg taaggaggga	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct	caggagagga ggggagaatt gctgaatggc cctgcctgca gcagcagctg	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac	gccacttaac tgtccctgcc cccagccctg ccccaccctc ctaggccagg	480 540
cctagtcacc t cttctgacct c caggctccac t gtggggccgc c	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc	caggagagga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga	gccacttaac tgtccctgcc cccagccctg cccaccctc ctaggccagg gccctgcctt	480 540 600 660
cctagtcacc to cttctgacct caggctccac to gtggggccgc ctcagttgcca a	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg	caggagagga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc	gccacttaac tgtccctgcc cccagccctg ccccaccctc ctaggccagg gccctgcctt tggcagccc	480 540 600 660 720
cctagtcacc to cttctgacct of caggctccac to gtggggccgc of tcagttgcca atggggactgg	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgcccc	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga	caggagaga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg	gccacttaac tgtccctgcc cccagccctg cccaccctc ctaggccagg gccctgcctt tggcagcccc ttgggcccgg	480 540 600 660 720 780
cctagtcacc to cttctgacct caggctccac to gtggggccgc ctcagttgcca a	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgcccc	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga	caggagaga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg	gccacttaac tgtccctgcc cccagccctg cccaccctc ctaggccagg gccctgcctt tggcagcccc ttgggcccgg	480 540 600 660 720 780 840
cctagtcacc to cttctgacct of caggctccac to gtggggccgc of tcagttgcca atggggactgg	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgcccc	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga	caggagaga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg	gccacttaac tgtccctgcc cccagcctg cccacctc ctaggccagg gccttgctt tggcagcccc ttgggcccgg	480 540 600 660 720 780 840
cctagtcacc to cttctgacct of caggctccac to gtggggccgc of tcagttgcca atggggactgg	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgcccc	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga	caggagaga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg	gccacttaac tgtccctgcc cccagcctg cccacctc ctaggccagg gccttgctt tggcagcccc ttgggcccgg	480 540 600 660 720 780 840
cctagtcacc to cttctgacct of caggetccac to gtggggccgc of teagttgcca at tggggactgg of ageageatet to constitute the constitute of the constit	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgcccc	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga	caggagaga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg	gccacttaac tgtccctgcc cccagcctg cccacctc ctaggccagg gccttgctt tggcagcccc ttgggcccgg	480 540 600 660 720 780 840
cctagtcacc to cttctgacct of caggetccac to gtggggccgc of tcagttgcca at tggggactgg of agcagcatct to cagge caggactct of caggactct of cagge caggactct of caggactca caggactca caggactca caggactca caggactct of caggactca caggactc	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgcccc	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga	caggagaga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg	gccacttaac tgtccctgcc cccagcctg cccacctc ctaggccagg gccttgctt tggcagcccc ttgggcccgg	480 540 600 660 720 780 840
cctagtcacc to cttctgacct of caggetccac to gtggggccgc of tcagttgcca at tggggactgg of agcagcatct to cagge caggactgg of cagge caggactgg of cagg	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgcccc tccggtgcta	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga	caggagaga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg	gccacttaac tgtccctgcc cccagcctg cccacctc ctaggccagg gccttgctt tggcagcccc ttgggcccgg	480 540 600 660 720 780 840
cctagtcacc to cttctgacct gaggtccac to gtggggccgc tcagttgca at tggggactgg agcagcatct to compare to the compare t	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgcccc tccggtgcta	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga	caggagaga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg	gccacttaac tgtccctgcc cccagcctg cccacctc ctaggccagg gccttgctt tggcagcccc ttgggcccgg	480 540 600 660 720 780 840
cctagtcacc to cttctgacct of caggetcac to gtggggccgc of tcagttgca at tggggactgg agcagcatct to compare to the com	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgccc tccggtgcta	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga tcctccctc	caggagagga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc ccaccctca	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg cagctcaagc	gccacttaac tgtccctgcc cccagcctg cccacctc ctaggccagg gccctgcctt tggcagcccc ttgggcccgg caagtccag	480 540 600 660 720 780 840 899
cctagtcacc to cttctgacct gaggtccac to gtggggccgc tcagttgca at tggggactgg agcagcatct to construct the construction of construction construct to construct the construction of construction construct to construct the construction of construction constructi	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgccc tccggtgcta	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga tcctccctc	caggagagga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc ccaccctca	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg cagctcaagc	gccacttaac tgtccctgcc cccagccctg cccaccctc ctaggccagg gccctgcctt tggcagcccc ttgggcccgg caagtccag	480 540 600 660 720 780 840 899
cctagtcacc to cttctgacct of caggetcac to gtggggccgc of tcagttgca at tggggactgg agcagcatct to compare to the com	gtgacacagt tagctgtgag gtggctctgg taaggaggga cgcagatggg aaagctgcat gcgctgccc tccggtgcta	tgttcagagt ggtgcctgag agcccctgcc ggggctgtct ctcaggaagc caggggaatg tgggaggga tcctccctc	caggagagga ggggagaatt gctgaatggc cctgcctgca gcagcagctg cccaggtgca cggcaaggca gagcctggcc ccaccctca	tgggcgcact caccccattc tccaccctcc tccccgagca cactcagcac ctcagcgaga cacagggctc agggctggtg cagctcaagc	gccacttaac tgtccctgcc cccagccctg cccaccctc ctaggccagg gccctgcctt tggcagcccc ttgggcccgg caagtccag	480 540 600 660 720 780 840 899

aaaaaaaa

aactctactt	ctagcttctt	gcttctgggc	tgctgctata	ccaaacagga	atgtaatact	180
ttctgtcagc	ttcaggcctt	tgcacatgca	gttcactttg	tctatcttgg	tttttattct	240
taggatttta	attctcctaa	gaagctttct	ctgaccagcc	taaaacttac	gtaagccctg	300
ggttaggtgc	tatgcttatg	tcctcccata	gcattttgca	tttgcatgtg	ttgtaactct	360
taatgtacag	catcatgatt	gcctatttta	actttcctgt	ttgttacagt	agactttaat	420
ctctttaagg	acaggaactg	tgtcttgttt	agaatcccca	gagcttattt	agtacaatgg	480
ctatgcttat	aatttaagta	tttattgaac	aaatgaaatt	ttcctaagcc	ctaaaacctt	540
gcaagatgtt	ttagtgcagg	aaactggcct	cggtggagtt	gaataactag	cacgaggtca	600
ctcacctaaa	aagtggtgag	gagggattaa	aatctaaatc	tgtttagctg	taaagattgg	660
gcttttttyc	ttgctgctgc	acatgactgc	yctctctcat	gttgcctgta	cacatccctg	720
tcaagtgttc	aaacagcccg	tgcctaacaa	ccccatccat	agcttctgag	gaaagttgtg	780
tcatctttgg	acagctctga	gagctgaagc	gagtctttgc	agaataattt	cccatctatt	840
ggtcttaatt	tatgctttgg	agaatataac	ttattttcaa	aaaacaaatg	attcagaatt	900
tgtcatctcc	ttaaggtccg	tttattagtt	tatttcattc	cttcattcac	tgataaccat	960
ttactgagca	ccagcctggg	caacatggtg	agaacccatc	tctaccaatt	taaaaaaaaa	1020
aaaaaagggc	ggccgctcta	gaggatccaa	gcttacgtac	gcgtgcatgc	gacgtcata	1079
<210> 62						
<211> 1928						
<212> DNA						
<213> Homo	sapiens					•
400 60						
<400> 62			h	~~+~~~~~	tagastataa	60
ggcacgagag	taggtctgcc	ggcgatggag	tggtgggeta	getegeeget	ttcaaaatgg	120
ctgctgttgt	tcctcctgcc	ctcagcgcag	ggccgccaga	aggagtcagg	ttcaactcaa	180
aaagtattta	ttgaccaaat	taacaggtet	reggagaatt	acquactacq	accaaccatc	240
aactgcagct	gctaccatgg	tgtcatagaa	gaggatetaa	eccectecte	toagateact	300
tccaggaaga	tgatggcaga	ggtagtcaga	cggaagctag	ggacccacca	tagattaga	360
aagaacagac	tgtaccggga	aaatgactgc	atgitteett	caaggegeag	castatacas	420
cactttattt	tggaagtgat	egggegtete	cetgacatgg	cactettete	cttcagtaag	480
gattateete	aggttcctaa	atgyatygay	gettggagat	tttaaaaaa	gggacctgct	540
acatcagagt	accatgatat	catgtateet	castagasee	tetteagaga	agatctggta	600
gtttggccaa	tttatcctac	aggictigga	aaaaactcta	caccatattt	ccaaggatca	660
aggtcagcag	cacagtggcc cagaacgaga	atgyadadag	attatatata	ggaaaaaccc	aaaacttott	720
aggacaagtc	acaccaaaaa	ceececcact	asstotatos	aaratacctt	addaaaagca	780
gatgcagaat	atgtccatct	tatagetcec	tacaaataca	agtatctgtt	taattttcga	840
gccgccaagg	caagtttccg	atttaaacac	ctcttcctat	gtggctcact	tottttccat	900
attactasta	agtggctaga	attettetat	ccacagetga	agccatgggt	tcactatatc	960
ccactcaaaa	cagatetete	caatgtccaa	gagctgttac	aatttotaaa	agcaaatgat	1020
gatgtagete	aagagattgc	tgaaaggga	agccagttta	ttaggaacca	tttgcagatg	1080
gatgacatca	cctgttactg	ggagaacctc	ttgagtgaat	actctaaatt	cctgtcttat	1140
aatgtaacga	gaaggaaagg	ttatgatcaa	attattccca	aaatgttgaa	aactgaacta	1200
tagtagtcat	cataggacca	tagtcctctt	tgtggcaaca	gatctcagat	atcctacggt	1260
gagaagetta	ccataagctt	ggcacctata	ccttgaatat	ctgctatcaa	gccaaatacc	1320
tagttttcct	tatcatgctg	cacccagage	aactcttgag	aaagatttaa	aatgtgtcta	1380
atacactgat	atgaagcagt	tcaacttttt	ggatgaataa	ggaccagaaa	tcgtgagatg	1440
tggattttga	acccaactct	acctttcatt	ttcttaagac	caatcacage	ttgtgcctca	1500
gatcatccac	ctgtgtgagt	ccatcactgt	gaaattgact	gtgtccatgt	gatgatgccc	1560
tttgtcccat	tatttggagc	agaaaattcg	tcatttggaa	gtagtacaac	tcattgctgg	1620
aattotoaaa	ttattcaagg	cgtgatctct	gtcactttat	tttaatgtag	gaaaccctat	1680
ggggtttato	aaaaatactt	ggggatcatt	ctctgaatgg	tctaaggaag	cggtagccat	1740
gccatgcaat	gatgtaggag	ttctcttttg	taaaaccata	aactctgtta	ctcaggaggt	1800
ttctataato	ccacatagaa	agaggccaat	tgcatgagta	attattgcaa	ttggatttca	1860
ggttcccttt	ttgtgccttc	atgccctact	tcttaatgcc	tctctaaagc	caaaaaaaa	1920
2222222		-	_			1928

```
<210> 63
<211> 781
<212> DNA
<213> Homo sapiens
<400> 63
ggcacgagat tttcagcctt tttggactgg tttctccaca tcttcgtgga tttatctaac
                                                                       60
tttggtcttt gatgttggtg accttcagat tgggtctctg agtgaacatc ctttttgttg
                                                                      120
atgttgatac tattcctttc tgtttgtttg tttgttttcc ttctaacagt cagggccctc
                                                                      180
tgctgcaggt ctgctggagt ttggttgagg tccactccag accetgtttg tctgggtttt
                                                                      240
gccagaggag gctgcagaat agcaatgatt gctgcctgtt tttcctctgg aagctttgtc
                                                                      300
ccagaggggc acccaccaga tgccagccag agctctcctg tatgaggtgt ctgttggccc
                                                                      360
                                                                      420
atacttggag gtgccttcca gtcaggatac acaggtgtca ggtacccact tgaggaggca
                                                                      480
ctctgtcccc tatcagagct cgaacactgt gctgggagat ccactgttct cttcagagct
gtcagacagg gacgtttaag tctgctgaag ctatgcccac agctgcccct ttccccagat
                                                                      540
                                                                      600
gctctgtccc agggagaagg gagttttatc tataagtctc tgactggggc tgctgccttt
                                                                      660
tcttcagaga tgccctgccc caagacgggg actctagaga ggcagtctgg ctgcagtggc
cttgctgaac tgtggtgggc ttcacccagt tggaccttcc ctgagccttt ttttttaccc
                                                                      720
tgtgagggta aaaatgccta atcaagcctc agcaatggtg gatgcccttc cccccaccaa
                                                                      780
                                                                      781
g
<210> 64
<211> 1194
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1172)
<223> n equals a,t,g, or c
<400> 64
                                                                       60
ggcacgagaa gacatggagt cttaagtgtg atcagtggga gggggctgga atcatttaga
ggcatcttca ttcacaaaac caggagctga tactggctgt cagccaggac ttcaactgac
                                                                      120
                                                                      180
ctatgtagaa cctgtccatg tggcccctcc ttgcagtctc cccatttggg ctggtttggg
                                                                      240
cttcatcaca gtccggcagc ttacttctaa gggcaagcat tccacgacaa cacagcagaa
                                                                      300
gggcatggca tttttacagt gaagtttggc aatctcatag cgtcgcttct gtcctacttt
                                                                       360
420
cgaccacgcg atggaagaag tgacacggtc atgttatgag aggagtgtgt gggatgggag
                                                                       480
atagggctgt ggccacctgc agaaaacagc atctgctata ggctgtcatg gaagcgcagg
atggggattt agcctacctg aggggtcagt cagcaaaggc ctctgggagg aagtgagatc
                                                                       540
                                                                       600
ttcggctgag gatgtgaagg gctaaaagga gaatgaggaa gagtttcagg gagaggaatc
                                                                       660
aatgaaacga gtccagagac gctggtgagt tggatggttt gcttcagtat gatgacaata
                                                                       720
cagaggggca aggagactgg tgcaggagaa gagagaaggt gccatgtgct ctgggtcgtg
                                                                       780
tcttctatgc cagactccct tagaagagga gcagcctcca gtcagcggtg tcccaggaac
                                                                       840
acggaggcta gacaggacaa tggcagccaa tccctgctcc caaactggtg acagtgggga
                                                                       900
aaagctgcat ggtctagatc caccctgctc cctggcccca gtatagaaga tcaaattcaa
                                                                       960
tctgcccaat cttatccaga taaagtaaag gaagactgga aaaaagaact aatccacggc
                                                                      1020
tccatctgcc catgactttc tctgctgatg ccggaggcag ctatggataa agagacggca
cacggcatgt cccgacgctg tggaggtggg gagaccccgc aagtccacag gaaaagagtt
                                                                      1080
aagttgctgc cacctgggca tccgctattc tctgctcttc tgcctcatcc tcaattcaga
                                                                     1140
                                                                      1194
ccatgatgga gctgattgtc tcccatttta tnccttggat tgaatggtct cgag
```

<210> 65 <211> 1677

```
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1012)
<223> n equals a,t,g, or c
<400> 65
                                                                        60
ggtgcagtgg tgccatcaca gttcactgca gccttgacct cccgggctca agcaatcctc
ccacctcage cacttgagta getgagacet cagatatgtg ccatcacace cagetgattt
                                                                       120
tttaaaatta attttttgta gagatagggt ctcatatgtt gcccatgctg gtctcaaact
                                                                       180
actgggttca aatgatcctc ctgcctcagc cttccaaagt actgggatta caggcatgag
                                                                       240
ccaccatgcc gggctgggag gcggaatttt gttcagtcta aagataagct ttttcatagc
                                                                       300
                                                                       360
tctggctgta gtgggaggga gcagaggagt gaatgattgt cagttgggag ggtgcagagt
                                                                       420
gggctcctgc cctagggtgr aggtragggt ggcttaggtg asmcamcaca gaggcctgt
teagececae gteecetece tgtgeteect cetectete ceteteetge aggegtggra
                                                                       480
                                                                       540
ggtatcatca ttcagcagat ttcaccagag gcagtggagg aggcaggtac ctgagccaga
                                                                       600
attcagaatg tottattoto cacttgacto tgccactaac ttgttgtgca actttgggco
                                                                       660
tttccccagg ccttcatttt cttttctttt ctttttcttt ytttttttt gaggcggagt
ctcgctatgt tgcccaggct ggagtgcagt ggcgcagcat catctcggct cactgcaagc
                                                                       720
tocaccttct gagttcacgc cattctactg cctcagcctc ccgagtagcc gggactgcag
                                                                       780
gcrcccacca ccacgcccgg cttattttt gtatttttag tagagacagg gtttcaccac
                                                                       840
                                                                       900
gttagccaag atggtctcga tctcctgacc tcgtgatcca cccgcctggg cttcccaaag
                                                                       960
tgctgggatt acaggcgtga gccactgcgc ccggccattt tcttaaatat ctaataaaaa
atatatagca aatgcagttt ttaaactacg acaatatgac cacgcaaaag antattatct
                                                                       1020
tccaagactg ctggtccaag gaaaagtcag taataaagtg gaagcattgt agcttatgga
                                                                       1080
atgactggtt asatttggga gaagccttag caataatcta gaatctgcat agataataca
                                                                       1140
tctgaggatt gggctttgtg gtttacaaag cattttttt tcctcttttg atcccagccg
                                                                       1200
tttgtctgga ctgatacaaa gcatttttat tagtttgtct tattcaatcc tcacaccacc
                                                                       1260
                                                                       1320
tcaaatttac agaggatatg gatctggtta atttgtatga ctatgtaacc tcatgtcagt
                                                                       1380
ccacagcact gcctggaggt gggtagaggt ggtcctgggc tggaatccca gccccagtgg
                                                                       1440
gaccttgagc aagttacttt agctgtctgc acctaaattt cctcactggc aaaacaggaa
tactggtggt tcacacctgc aattccagca ctttgggagg ctgaggtggg aggattgctt
                                                                       1500
gagtccagaa gttcaaaacc agactgggca acatagcaag accatctcta caaaaattaa
                                                                       1560
ataaataaaa catttacaag ggttgtggtg aagattaaat gagatcactc acgaaaaagc
                                                                       1620
tcagcagacc ctgatgtgca gtaggtgctc aataaatgtt agccagcaaa aaaaaag
                                                                       1677
<210> 66
<211> 1237
<212> DNA
<213> Homo sapiens
<400> 66
                                                                         60
agcaaaccca ggaaggtgtg gcgtccccgc ttcgcgccaa gatggtgctg gtgctgcgcc
                                                                        120
atcetttgtg tgcccgggaa agggcgttcc gggagccggg tcgggggctc ctgactcgca
ctgggcagca tgacggtgcg ccggctgtca ctgctgtgcc gggacctctg ggcgctgtgg
                                                                        180
                                                                        240
ctgctgctga aggccggcgc agtgcgtggg gcgcgggcgg gtcctcgcct ccccggaagg
                                                                        300
tgttgtgggg cgacatgcgg ggacgccggg cgggggtgga cgttctgggc ccagccctgt
cctcagaagc tgctggggca gaagcccggg gctgggggat gccggggatg ggtgttgggg
                                                                        360
                                                                        420
tgggtgcctc cgagaccaga ggagccctgt tccttggcag ggaaggtgtg cacgggcctt
                                                                        480
gcccgatgga tggtttaggg ccatggccct ggggtccctg gtgagcagtg gggccgcctc
                                                                        540
tgcccttggc ctgtgaggga ctgtctgtgc tggtcccaga aggctgggat cacctttcca
ctggctcctt tgttcgaggt ttttcataga caggctatgt ggacaaatga gggcagcgcc
                                                                        600
cacgtctggc tggtggaggg gctgcggctc ctccttggag gggacgcctg gccactgctg
                                                                        660
                                                                        720
tccccacaat ggggccaccc gtggtgcaag gcgtgacaag ctgccctctc taggtaagca
ggacttggga ggcccctggc caagcctgtg gacccggctg ggcggcctct gtggtctcag
                                                                        780
```

```
gtttgggtgt gtttggtctg gtcagggctc aggggctgct ggtccacact ggccccatcc
                                                                       840
 tgacaattgg agctttgggg caaggtccct ggagaagggg tcacgtcggg aggaaacagc
                                                                       900
 ctgggttttg ttgatgcttt tctaagaatg gagtactcgt tttcaagaga tttgtcctaa
                                                                       960
 ttatattttc cagcgggtac ttatgccaag tattgatgaa taattcataa aataagcatc
                                                                      1020
 tttgtgaatt ttagtgaatc agacettaae tateaaegge aatgaatgaa eatetaaagt .
                                                                      1080
 ttccaatttt aaagtaaaga actggctggg tacagcagtt cacgcctgta atcccagcac
                                                                      1140
 tttgggaggc caaggctaga ggatcgcttg agcccaggag tttgagatca gcctgggcaa
                                                                      1200.
 cataccaaga cctcatctgt taaaaaaaa aaaaaaa
                                                                      1237
 <210> 67
 <211> 1934
<212> DNA
<213> Homo sapiens
<400> 67
ccacgcgtcc ggggcgttcc tggtcgtgag aggggagccc caggggagct ggggcagcat
                                                                        60
gactggggtg ataaatggcc ggaaatttgg cgtggccaca ctcaacacca gcgtgatgca
                                                                       120
ggaggcacac tccggggtca gcagcatcca cagcagcatc cgccatgtcc cagcaaacgt
                                                                       180
ggggcctctg atgcgggtgc tcgtggtcac catcgccccc atctactggg ccctggccag
                                                                       240
agagagtggg gaagccctga atggccactc tctgactggg ggcaagttcc ggcaggagtc
                                                                       300
acacgtggag tttgctacag gggagctgct cacgatgacc cagtggcccg gggtctggat
                                                                       360
cccgatggcc tcctgctcct cgacgtggtg gtcaatggcg ttgtccccgg acagcctggc
                                                                       420
tgacgcagat cttcaagtgc aggactttga ggagcactac gtgcaaacag ggcctggcca
                                                                       480
getgttegtg ggetecacae agegettett ceagggegge eteceetegt teetaegetg
                                                                       540·
caaccacage atccagtaca acgeggeeeg gggeeeecag eeceagetgg tgcageacet
                                                                       600
gcgggcctca gctatcagct cggcctttga tccagaggcc gaggccctgc gcttccagct
                                                                       660
cgctacagcc ctgcaggcgg aggagaacga ggtcggctgc cccgagggct ttgagctgga
                                                                       720
ctcccaggga gcgttttgtg tggatgtgga cgagtgtgcg tgggatgctc acctctgccg
                                                                       780
agagggacag cgctgtgtga acctgctcgg gtcctaccgc tgcctccccg actgtgggcc
                                                                      840
tggcttccgg gtggctgatg gggccggctg tgaagatgtg gacgaatgcc tggaggggtt
                                                                      900
ggacgactgt cactacaacc agctctgcga gaacacccca ggcggtcacc gctgcagctg
                                                                      960
1020
gcagctgccc aaggcctgcg cctaccagtg ccacaacctc cagggcagct accgctgcct
                                                                     1080
gtgccccca ggccagaccc tccttcgcga cggcaaggcc tgcacctcac tggagcggaa
                                                                     1140
tggacaaaat gtgaccaccg tcagccaccg aggccctcta ttgccctggc tgcggccctg
                                                                     1200
ggcctcgatc cccggtacct cctaccacgc ctgggtctct ctccgtccgg gtcccatggc
                                                                     1260
cctgagcagt gtgggccggg cctggtgccc tcctggtttc atcaggcaga acggagtctg
                                                                     1320
cacagacett gacgagtgee gegtgaggaa eetgtgteag caegeetgee geaacaetga
                                                                     1380
gggcagctac cagtgcctgt gccccgccgg ctaccgtctg ctccccagcg ggaagaactg
                                                                     1440
ccaggacatc aacgagtgcg aggaggagag catcgagtgt ggacccggcc agatgtgctt
                                                                     1500
caacacccgt ggcagctacc agtgtgtgga cacaccctgt cctgccacct accggcaggg
                                                                     1560
ccccagccct gggacgtgct tccggcgctg ctcgcaggac tgcggcacgg gcggcccttc
                                                                     1620
tacgctgcag taccggctgc tgccgctgcc cctgggcgtg cgcgcccacc acgacgtggc
                                                                     1680
ccgcctcacc gccttctccg aggtcggcgt ccccgccaac cgcaccgagc tcagcatgct
                                                                     1740
ggagcccgac ccccgcagcc ccttcgcgct gcgtccgctg cgcgcgggcc ttggcgcggt
                                                                     1800
ctacacccgt cgcgcgctca cccgcgccgg cctctaccgg ctcaccgtgc gtgctgcggc
                                                                     1860
accgcgccac caaagcgtct tcgtcttgct catcgccgtg tccccctacc cctactaaac
                                                                     1920
gggagagggc attg
                                                                     1934
<210> 68
```

<211> 3300

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

```
<222> (1)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (3)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (15)
<223> n equals a,t,g, or c
<400> 68
nengeageeg gaegneegag egeagegagt cagtgagega ggaageggaa gagegeecaa
                                                                       60
tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg cacgacaggt
                                                                      120
ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag ctcactcatt
                                                                      180
aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga attgtgagcg
                                                                       240
gataacaatt tcacacagga aacagctatg accatgatta cgccaagctc gaaattaacc
                                                                       300
ctcactaaag ggaacaaaag ctggagctcc accgcggtgg cggccgctct agaactagtg
                                                                       360
gatcccccgg gctgcaggaa ttcggcacga gaacacatct taagggaacc aagtctcaag
                                                                       420
agaaatcaag taattatgaa tgaacagctc taaaaaagag agagagaata ttttcttaaa
                                                                       480
tcaacttagt tgctgttatg accaaagaac agatgttgtg gtgttcaccc cagagaagca
                                                                       540
agagattttc ccttaaacct cagcttataa tgaatggaag tgaatgacag ggagagagtt
                                                                       600
tttctctcgt ttcccagaac tctatccttt tcttctcaaa cagttggaaa ctgtagccaa
                                                                       660
tacagtagac agtgatatgg gagaaccaaa tcgtcatcca agcatgtttc tcttactttt
                                                                       720
ggtgttggag agactctacg cttccccgat ggatggtact tcttctgctc tcagcatggg
                                                                       780
accttttgtt cccttcatta tgaggtgtgg tcactcacct gtctaccact cccgtgaaat
                                                                       840
ggcagctcgt gccttggtcc catttgttat gatagatcac attcctaata ccattcgaac
                                                                       900
tetgttgtcc acactececa getgcaetga ceagtgttte eggcaaaace acatteatgg
                                                                       960
gacacttete caggittite attigitgea agectactea gactecaaae aeggaaegaa
                                                                      1020
ttcagacttc cagcacgagc tgactgacat cactgtttgt accaaagcca aactctggct
                                                                      1080
ggccaagagg caaaatccat gtttggtgac cagagctgta tatattgata ttctcttcct
                                                                      1140
attgacttgc tgcctcaaca gatctgcaaa ggacaaccag ccagttctgg agagtcttgg
                                                                      1200
cttctgggag gaagtcagag ggattatctc aggatcagag ctgataacgg gattcccttg
                                                                      1260
ggccttcaag gtgccaggcc tgccccagta cctccagagc ctcaccagac tagccattgc
                                                                      1320
tgcagtgtgg gccgcggcag ccaagagtgg agagcgggag acgaatgtcc ccatctcttt
                                                                      1380
ctctcagctg ttagaatctg ccttccctga agtgcgctca ctaacactgg aagccctctt
                                                                      1440
                                                                      1500
ggaaaagttc ttagcagcag cctctggact tggagagaag ggcgtgccac ccttgctgtg
caacatggga gagaagttct tattgttggc catgaaggaa aatcacccag aatgcttctg
                                                                      1560
caagatactg aaaattctcc actgcatgga ccctggtgag tggcttcccc agacggagca
                                                                      1620
ctgtgtccat ctgaccccaa aggagttctt gatctggacg atggatattg cttccaatga
                                                                      1680
aagatctgaa attcagagtg tagctctgag acttgcttcc aaagtcattt cccaccacat
                                                                      1740
gcagacatgt gtggagaaca gggaattgat agctgctgag ctgaagcagt gggttcagct
                                                                      1800
ggtcatcttg tcatgtgaag accatcttcc tacagagtct aggctggccg tcgttgaagt
                                                                      1860
cctcaccagt actacaccac ttttcctcac caacccccat cctattcttg agttgcagga
                                                                      1920
 tacacttgct ctctggaagt gtgtccttac ccttctgcag agtgaggagc aagctgttag
                                                                      1980
 agatgcagcc acggaaaccg tgacaactgc catgtcacaa gaaaatacct gccagtcaac
                                                                      2040
 agagtttgcc ttctgccagg tggatgcctc catcgctctg gccctggccc tggccgtcct
                                                                      2100
 gtgtgatctg ctccagcagt gggaccagtt ggcccctgga ctgcccatcc tgctgggatg
                                                                      2160
 gctgttggga gagagtgatg acctcgtggc ctgtgtggag agcatgcatc aggtggaaga
                                                                      2220
 agactacctg tttgaaaaag cagaagtcaa cttttgggcc gagaccctga tctttgtgaa
                                                                      2280
 atacctctgc aagcacctct tctgtctcct ctcaaagtcc ggctggcgtc ccccaagccc
                                                                      2340
 tgagatgctc tgtcaccttc aaaggatggt gtcagagcag tgccacctcc tgtctcagtt
                                                                       2400
 cttcagagag cttccaccag ctgctgagtt tgtgaagaca gtggagttca caagactacg
                                                                       2460
 2520
 aggggaagac accetagtte teagtgtttg ggaetettat geagaatega ggeagttaae
                                                                       2580
 tcttccaaga acagaagcgg catgttgaag aaaatctggg ggattgggat gggggtatgt
                                                                       2640
```

<400> 70

gtggatttt	cctccactaa	atctgcagga	aacatgttga	acataaattc	aaaaatttta	2700
tcccaaaaaa	aaaaaaaaa	aaactcgagg	gggggcccgg	tacccaattc	gccctatagt	2760
gagtcgtatt	acaattcact	ggccgtcgtt	ttacaacgtc	gtgactggga	aaaccctggc	2820
gttacccaac	ttaatcqcct	tgcagcacat	cccctttcg	ccagctggcg	taatagcgaa	2880
gaggcccgca	ccgatcgccc	ttcccaacag	ttgcgcagcc	tgaatggcga	atggcaaatt	2940
gaggacagta	atattttgtt	aaaattcqcq	ttaaattttt	gttaaatcag	ctcattttt	3000
aaccaatagg	ccgaaatcgg	caaaatccct	tataaatcaa	aagaatagac	cgagataggg	3060
transtatta	ttccagtttg	gaacaagagt	ccactattaa	agaacgtgga	ctccaacgtc	3120
cigagigicig	aaaccgtcta	tcaggggat	ggcccactac	gtgaaccatc	accctaatca	3180
adayyycydd	ggtcgaggtg	ccataaaaca	ctaaatcgga	accctaaagg	gagcccccga	3240
agtititity	gacggggaaa	accacaaac	gtggcgagaa	aggaagggaa	gctgtctctt	3300
tttagagett	gacggggaaa	geeggegaae	509905-5			
<210> 69						
<211> 1797	•					
<212> DNA						
<213> Homo	canione					
(213) HOMO	saprens		•			
<400> 69					•	
autcascaut	atcgataagc	ttgatatcga	attcctgcaa	cagttcttgg	aaacccactc	60
ggccgacggc	cgcctccatt	caccaggcca	cocatcacaa	gaggcaacac	caggagccaa	120
catgagggcca	gggactgaac	tactataacc	cagagcagcg	ctgctggtgc	tgttgggggt	180
catgageteg	ctgtgtgtgc	actactcaca	cccaggtgca	aagaggtcag	agaaaatcta	240
ggcagccagc	agtctgcgtg	accaccaca	gagetttacg	gggtcccgga	cctactcctt	300
ccagcagaga	gcatggccag	aggaccaaca	gagestages	cccacaagga	aggacaagct	360
ggtcgggcag	geatggeeag	tagagatee	aggatatettee	aggtaccaga	acttcagcaa	420
gttgcaattc	taccccagcc	tggaggattt	catacacccc	artgccatgg	agtattacaa	480
aggaagcaga	cacgggtcgg	aggaageeta	tacagacccc	aattectacq	agaatgtgct	540
ctgggggcgg	ttctcgaagc	ccccagaaga	taccesacea	garggratag	ataaceteta	600
catttgcaag	cagaaaacca	cagagacagg	rgcccagcag	ccacttcta	atctctatcc	660
cagaggggac	ctcagcctgt	cactggccct	gaagactggc	tragaacttc	agcattccat	720
ctctgcctcc	ccggaagaag	atgaaggaat	tegaggatta	ccagaacece	aagcatcccc	780
ccattcagtg	gcgcgagtcc	aggaaggtca	tggggcaact	ccagagaaga	atggggaggt	840
tggcccggtg	ggaagcccag	acgaggagga	cggggaaccy	gattacgtga	argaggagge	900
ggcagccaca	gaagcctagg	gcagaccaag	aagaaaggag	ccaaggcaaa	taaactttta	960
tgtgctcatg	gacccatcgc	tgccttccaa	ggaccatttc	ccagagetac	ccaacccca	1020
agcccctgcc	atggttgctc	ctggaaggag	aaccagccac	cctgaggacc	tooggecat	1080
gcgtgcacag	cctgggaaaa	gacagttact	cacgggagct	geaggeeeeg	ccaccaagee	1140
ctctcccgac	ccaggctttg	tggggcaggc	acctggtacc	aagggtaacc	cggctcctgg	1200
tatggacgga	tgcgcaggat	ttaggataag	ctgtcaccca	gtccccataa	caaaaccact	1260
gtccaacact	ggtatctgtg	ttcttttgtg	ctatgaattt	ggattcctaa	ttgctattgt	
tggttgctgg	ggttttaaat	gattgataag	cttgtacagt	taacttatag	agggggagcc	1320
atatttaaca	ttctggattt	cagagtagag	atttctgtgt	. tgtctcctag	aaagcattac	1380
atgtagttta	tttcagcatc	cttgttgggt	ggggccctgg	ctctcttccc	ctttggtggg	1440
acctcccctt	tetttagget	tcagttcact	caggaagaaa	tgaggctgtc	gccatcttta	1500
tatacttcca	gtggaaatgt	cacttgctac	agacaatagt	gcatgagagt	ctagagaagt	1560
agtgaccaga	acagggcaga	gtaggtcccc	tccatggccc	: tgaatcctcc	tctgctccag	1620
gactageete	tocagageto	attaaacagt	gttgtgactg	, tctcatggga	agagctgggg	1680
cccagaggga	ccttgagtca	gaaatgttgc	cagaaaaagt	: atctcctcca	accaaaacat	1740
ctcaataaaa	ccattttagt	tgaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaa	1797
	_					
<210> 70						
<211> 1373	3					
<212> DNA						
<213> Homo	sapiens					
	-					

ggcacgaggg ctgacggcgc ttttgtctcc ggtgagtttt gtggcgggaa gcttctgcgc

⁶⁰

		tectacacaa	cctgctccta	cagccggcga	120 [.]
tggtgcttag taaccga	Ctt teeteeggae		agaaggagaa	ggcagctctg	180
tccactcccg gctgttc	ccc cggaggtcca	gaggeeeee	tatattaac	crctacctgt	240
tttctctgca gaggagt	agg gtcctttcag	ccatgaagca	cytyctyacc	ctagagggt	300
taggtgtggt actgacc	cta ctctccatct	tcgttagagt	gatggagete	accaacacac	360
tactagagag cccatcg	cct gggacctcct	ggaccaccag	aagecaacta	stasttast	420
ageccaccaa gggeett	cca gaccatccat	ccagaagcat	gtgataagac	CCCCCCCCC	480
actooccata ttttooa	aca ctgacctaga	catgtccaga	Lyggagteee	accesses	540
gacaagetga geacegt	tot aaccagagaa	ctattactag	gccccyaaaa	accegeeea	600
crogatoctc attocct	ggg caaggeetgt	ttaggccggt	Egeggeggee	Catgettgta	
atoctagoac titogga	age tgatgtgggt	ggatcacctg	aggicaggag	Cccagaccag	660
cotogcoac atggoga	aac cccatctcta	ctaaaaatac	aaaagccagc	cgggcgcgggc	720
agragagge tgtaate	cca gctccttggg	aggctgaggc	gggayaaccy	cccgaacccg	780
acceptage ttgcagt	nad ccdadatcqc	actgctgtac	ccagccuggg	Ccacagegea	840
agactccatc tcaaaaa	asa sagaaaagaa	aaagcctgtt	taatgcacag	gtgtgagtgg	900
attgcttatg gctatga	gat aggttgatct	caccettace	ccggggtctg	gtgtatgctg	960
tgctttcctc agcagta	taa etetaacate	tettagatgt	cccaacttca	gctgttggga	1020
gatggtgata ttttcas	togg coctgacate	acatctotct	gaggttcctt	tagtcttgaa	1080
tgtcttatgc tcaatta	the material	ctctcttcca	caagagetee	tccatgtttg	1140
tgtcttatgc tcaatta	att ggtgttgagt	attaggatga	gatggagtgt	tcagtgccca	1200
gatagcagtt gaagagt	grg rgggrgggc	geegggaega	gratattagt	ctgaagggg	1260
tttctcattt tacattt	taa agtegtteet	. ccaacatage	gegeeegge	ggcttaaatg	1320
tggtgggatg ccaaage	ctg ctcaagttat	. ggacattgtg	22222222	aaa	1373
attttttcta actaata	aaag tgggatatat	. atttaaaaaa	aaaaaaaaaa	444	

<210> 71 <211> 1579 <212> DNA

<213> Homo sapiens

<400> 71				actoacacoa	ctttactccq	60
ggcacgagga	tttggagggg	acaaacatcc	aaaccattta	agccacagca	tactctggaa	120
cagtgtgaat	aacacaggca	ttctcctaca	taatcacagt	acagecatea	caatgatatt	180
attgaatatc	atctaatata	ctttccatac	ccagattttc	ctayactecc	attacatttc	240
tcttactgtc	ctccctttag	ccttcctctt	tctccattca	ggattetate	ttaccacttt	300
attttcatgt	ctcttcagtc	tctctttagc	tttgttttc	tttcttgatg	ttcattttaa	360
taggaggcca	ggccagttgt	tttgtgaaag	atctgttctc	tttgatatgt	caccactatt	420
atttattta	ttatttttac	atgaatggat	tcaggctaaa	catttttggg	caggacagae	480
tattatatta	cctantgatg	tattettte	agtccatcat	ctggaggcac	Ctgatggtag	540
++++cccaat	attocoaaat	taagtttgat	tattttgtta	aggtagtgtt	Caccagaccc	600
atacatttta	aagacatcct	tttctctaat	tactcagtgg	actgtagagt	gatgettega	660
aactgaataa	ctaacactcc	ctaactcagt	gatttagcac	ccgttgattt	CCCCCCC	720
actosatosa	atattattat	agtagtttta	aatggtgatt	ttccatttct	accecege	780
taggtaggat	tettetataa	ttttatctt	atattttact	gggtgttaay	actigitation	840
cagettteet	regretttta	caccttttcc	cattctttta	tttttttcca	Coccee	
tectatttte	caatagatgg	atagaatttt	ctttctctgg	tttaaayytt	acacaccege	900
atatatatat	gtattctaaa	ccatttqccc	ttaaaacata	gagatggtta	Ciccigaciga	960
++=======	ctcagtaatg	ttactatctt	ctctttaata	agataggtac	LLLacticat	1020
tetatattet	taggtgaggt	ctccctcac	ccagtcaagt	tgatgttaat	Clayaattt	1080
tagttttta	aattatcgat	atactttctg	tcttttcctt	EEEEEEE	Colocodea	1140
tttatactt	ccccctttac	tctctttccc	ttcattcctt	CEEECECACE		1200
attttcact	cttaggtaga	ttatctttta	gaaaacagca	agatattata	taatttactt	1260
tatattetet	ttccaaatga	ttaaagtaat	aattaaaaat	ttttgatatg	tgtgtatgca	1320
acattagas	tectetteta	agragaagaa	ctctaccaca	tgcatgagtc	attagtgtgt	1380
aggataggaa	coccecege	taggtccago	tagatactct	gaagaaggta	ggtttcttca	1440
taaacactgg	gaagugguu	ccttacctta	aaaaaggagt	gaagaatact	gactgcagag	1500
gcccccatg	- claacigici	202020000	atagaactgg	aaaagaatat	taatatttgc	1560
		acacagaage				1579
caaaaaaaaa	aaaaaaaa					

```
<210> 72
<211> 1028
<212> DNA
<213> Homo sapiens
<400> 72
gcacgacaat tgaactgaac cctaaaaatg ctacttcaat tcaccttatg ggtatttggt
                                                                         60.
                                                                        120
gctatacatt tcccgaaatg ccttggtatc aaagaagaat tgctaaaatg ctgtttgcaa
                                                                        180
ctgcctccta gttccaccta tgagaaggta gtatgatgtc ctttgttaag ttagtacgga
                                                                        240
tttcttgaac cacagegeee attctaceat gtgttccaca cattgtggag ctctggattc
                                                                        300
agtgaagggg acttgaggca atttccttaa cgatccaatt caactgtgtt atcacaaggc
ttaacactta ttatccttga ctggtgagtg gttttctttt tccccgttag gtgagtggct
                                                                        360
ggtaattctg gaatactgtc atctaaaatg gctcgtggct aaaatctacc ttcattttct
                                                                        420
gtttgaaatc taaactatat tgaagtcata aaatagaaca agaaatacag catctgttac
                                                                        480
                                                                        540
ccagcatgtt ttagctgtat tacacacaat aacagaaaag taaagcagat gcttaagttg
                                                                        600
ataaaagaag aacactcatt ataacttcta ttttaaaaag catatgaaag gttcatattc
                                                                        660
tctcatattt tcaaggcctt ttgcttttct tgttaaaaat aagatttgag aggaatttct
                                                                        720
ggttaaactt tgggtttact catcacaagc ttttcagagt aagaaaacag gcaatcgaaa
aagctgtact tgtattattt acattataac aaggagcctt tttttctttc tgggaagcta
                                                                        780
tagtgtagaa attgatgtaa aaaatactta gttgtattct ttacacacag ttgagaaata
                                                                        840
                                                                        900
ttattaaaat aatgcaccaa tattttataa tggtattatt aaaataatgc ccatttgctg
gacacggtgg ctcatgcctg taatgccagc attttggaag gccaaggttg gtggatcagt
                                                                        960
                                                                       1020
tgagcctggg agtttgagac cagcctgggc aacgtggcaa aacccctgca aaaaaaaaa
                                                                       1028
aaaaaaa
<210> 73
<211> 3674
<212> DNA
<213> Homo sapiens
<400> 73
ggcacgagct caaaagaaat agggtgattt ttaaaggatt aataaaattc tgaaatgtta
                                                                         60
                                                                        120
agtagaagat tacattgtct agtcttgtat ttcctccttc tgttgctctc tttcattcac
acacteteag ttteteatat ttgtagetea tttatttggt tattteetaa gaatattgaa
                                                                        180
                                                                        240
agtgaagcaa ctatgtgact gtattcttca ggtaaacact gactgcgctt gttggatttt
ccctattttt gtgacttcaa gaataatatg ccctgctgaa tacatgccat ttcacattct
                                                                        300
gaaactgggt agagtggttg ggtgttctgc caacaattgc tagtggtgtg aattcattca
                                                                        360
                                                                        420
tatttgccag tattgctcac ttcaaagaaa ctccttcatc aagcagtcca gagctaggcc
                                                                        480
agatcaatgc tacaatcatg aagttctcat tgcatgcaat tgtgtaggat tgacaaggaa
ctcagataaa aatttccagg gtgcacttcc agaaccagct tcaacatatg tctacattgc
                                                                        540
ccccaagtta ataaagtgcc aaccctttac tctctcatac agccagaaat gttagaaatc
                                                                        600
caaaatcttg gtgcattatt ttttcataaa cgctaaaaca tttgaagaaa caatttaatt
                                                                        660
                                                                        720
atttaaaatt caagtatttt attcacatta tttgcaatat ccaaatgttt aaaaattccc
agataattaa ctagctatta cagatctcac ctagagggtt gatgttatga agactccagt
                                                                        780
                                                                        840
ggactgtact cacaaattga ctggacaccc tatgaaagtg ggtagacctc tcagcggaaa
ataagaaggg cttttaccta cagggcagga cagggtccca tgagagcagt tctgtggaga
                                                                        900
                                                                        960
tataaaaaga atggaagaag gaatgcctta tagtgatatt gtgacattat atctatatat
                                                                       1020
ctacatatat ctatctatct atatctacat ctatataatc ttacatttaa aattgtattc
ctacacatat tagaaactct tctaataaat gaagtaaaaa aattaaaaag aatacaaata
                                                                       1080
ttccagccc aaatgagaaa tcaaacatat taaaattgtt caagaaaatt tctttgaaca
                                                                       1140
cttctgaaag tttttggaaa cttagaaaag agggaaaaaa atccagtgtt actagtaatt
                                                                       1200
tccatggtaa tacagataaa atacattctt ttaattctgg gaaattagaa aaagtggggt
                                                                       1260
gatctttcca ggaaaaacat gtgtaacatc tgcttatcac tccagctccc tcctcctc
                                                                       1320
                                                                       1380
cctctccacg ttcccttgag taaatgtctg ggaaagcatg aagcttgatg caagaaccct
gttgtactgg cgttttcctc ccctgtgaaa acgtaactac tgttgggagt gaattgagga
                                                                       1440
tgtagaaagg tggtggaacc aaattgtggt caatggaaat aggagaatat ggttctcact
                                                                       1500
```

```
cttgagaaaa aaacctaaga ttagcccagg tagttgcctg taacttcagt ttttctgcct
                                                                         1560
  gggtttgata tagtttaggg ttggggttag attaagatct aaattacatc aggacaaaga
                                                                         1620
  gacagactat taactccaca gttaattaag gacgtatgtt ccatgtttat ttgttaaagc
                                                                         1680
  agtgtgaata gccttcaagc atgtgaataa tcttccatct tccccgccac acatacacac
                                                                         1740
  acacactttt tgtttctttc aggtagacac cttttaaaat gcagaactaa ctgaggcatt
                                                                         1800
  tcagtaactt tgctttcaaa tcaataaagt caaatgtatg gaaacatttt gtgccctact
                                                                        1860
  ctccataccc cgtgtactca aattctctac tgtatgaatt atgctttaag tagaattcag
                                                                        1920
  tgccaaggag aacttggtga aataaattat tttaattttt tttttatcct ttacaaagcc
                                                                        1980
 atggatttta tttggttgat gtgtgctctg tacacaagcc atttcaatag gatggagctg
                                                                        2040
 ttaattattt tccaaagagt aatagacatg caaaagtttc aataaaaact gggccattaa
                                                                        2100
 caagtaaatt aataaactaa taagcattcc cttctaggtt tttgccaaac tgcctatcca
                                                                        2160
 ataacaaatt tgagaatcgt tgtaaaagct agttatattt cagagaaatg attttcatta
                                                                        2220
 ttgaaactgt tctccctagc aggccatttt ccctttttcc tgggagttta gcaagtttag
                                                                        2280
 gagagaatag tcatgaaaag aaagggaaga aaggggagaa gggaagaggt taaaaagtaa
                                                                        2340
 gtgctcagac ctatgaacgt aatccctttg ctacaaatat ttaagagcag ctcagcttgg
                                                                        2400·
 ttgaaactga gttttgtcat cttccatatt tgcaggaagg tattttctga cttgcaatgc
                                                                        2460
 agctagatgt aaaattttat tttatcatcc tagaaagcct tgactagaaa aatgaataaa
                                                                        2520
 tattgagggt ttcctgtcca tatctggctt gcatgtgcca gaaagcagag aatagaaaat
                                                                        2580
 gtaatctcca acatccaagc atcgaaaccc aaggggtagg caattctatg taggttttgg
                                                                        2640
 acatgaagtt tggtgcatct tggtttatgc tggctcaact gctattaaac ctctctggct
                                                                        2700
 tatagtetet teattetatt agacaageae gtategaaca ettgettege acaaggetet
                                                                        2760
 ttagttaaca atttagcagc tactgtttgt gttaaacaca cttttcacca aataggttct
                                                                        2820
 gaggcaaacg agagcaatga ctatttaaag aaaggctttc ccagcatcac ttacacatcc
                                                                        2880
 caaaactaaa aagatcaact cttccaactg agaaaagact cctggctttg aatggaaact
                                                                        2940
 tacagcagag agtcacaggc cacggcaaca acaacgacaa caacaaacat ttggaatatt
                                                                        3000
 attctcaact cacgttttaa taatacatct tattattttt ctagtagaga aactacaaat
                                                                        3060
 cagcetette aacatttata tacagtttaa taageetett geaagttaet tgttetetea
                                                                        3120
 cctgaggtat tttttcctc cccaccttgc ccctgttcct cccttcctct tctccctttg
                                                                        3180
 caagaggaaa tatttaacat atttgggtcc aacttcaata atgtaataat taatacatta
                                                                       3240
 aaagcattta acttcctttc tagaaaaatg cacaggctaa ggcatagaca aaacaaagag
                                                                       3300
 aaatgctgag aaatttgcca ctggagacaa gcaatctgaa taaatatttg ccaaaagttc
                                                                       3360
 tttttatgtc atatagtgtc aggatttgaa ggagctattt ttttttaatg ttgcaactag
                                                                       3420
caactcatct tcggaagaca cagccaggag aatgaagtag aagtgaaagg tttataaatc
                                                                       3480
catttgtaag catttatccc atatatttta aattcaagaa aaattgtgtt tatctttaga
                                                                       3540
attttgtatt caatacttta tgtactatgt gactcatgct tctggataaa taaagcacca
                                                                       3600
aatatgtatc tgtaaccaca atcacacata ttatattaaa tatatatcta tataacaaaa
                                                                       3660
aaaaaaaaa aaaa
                                                                       3674
<210> 74
<211> 2797 .
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (853)
<223> n equals a,t,g, or c
<400> 74
ggcacgagag agcagacaga attatatgta gaggacacag gagatattta cattgtggat
                                                                         60
ggagatggag gattgaataa cagattgatc aaactgtccc aagatttcat gatcctttgg
                                                                        120
ctgcatggag aaaatgggac agggcctgct aagttcaaca tacctcacag tgttacactt
                                                                        180
gattcagctg gtcgggtgtg ggttgctgac cgaggaaata aaagaatcca agtatttgat
                                                                        240
aaagacactg gggagtggtt aggagcatgg aataattgtt tcacagaaga ggaccttctt
                                                                       300
cagtcagttt actcctgatg ggaagtactt gattgtggcc cagctgaatc ttagcaggct
                                                                       360
ctcagtcgta gcagcacccc cagtgggaag cattggggag tgttctgtga tcagcacaat
                                                                       420
ccaactagca gatcaagttt tgccacatct cctagaagtc gacagaaaag actggagcag
```

540 600

660

			45			
tctatgtagc	agaaattgga	gcaaaacaag	tacaaaaata	tatcccttta	aatagctatg	540
ttccttcart	taattcataa	tgtttctttc	ccgggaatat	ttcaagtggc	agttcagatt	600
		aaaatgatgt				660
aaagatctag	tatcagaaag	atttgtttt	gtatcattaa	gaatcttata	ttttgttgcc	720.
ctcttgggac	ttagttttat	ttgtaagtgc	ataaggatat	tttaatgaaa	ggaaagtaac	780
taaaaaatgg	ggttgggaag	agggactaag	gtggtaacct	cattatttgc	cctggtagac	840
tgattctccc	tonotaaaaa	aaatgggaat	aaaaatgagc	ttgcatgata	atttattaaa	900
		acctccagat				960
		ctattgatta				1020
		catgattgcc				1080
caagcaggca	aacagtcaca	acacagcaaa	agcgaccttg	gagcatagtg	ggacttttgg	1140
agtaggagtg	ctgcatttga	ctgagggaat	catggatact	tcgcaggaga	agtgaatttg	1200
agctcagact	tgaaaactga	ggaggagctt	accaagggac	aaggaggaga	aaacaataat	1260 [.]
		aaaagttaga				1320
gcctttttta	aagcccaact	tggcttctgt	ccattaccta	taagatattt	aatgtcagtc.	1380
agcctttaaa	tgtaggaata	aaatggctgg	catctaagca	ctttagtaaa	agaggtttt	1440
		gagcttcctt				1500
ttacatgaac	tcaacttatt	cctaacattt	gtctacctca	aagaaatttc	aagattattt	1560
		aaatcctttg				1620
tcctacatct	tctcctccca	ctccctcctt	tggtgtgaat	attggcttcc	caattaagac	1680
		tttagcttat				1740
taatctttca	aatcattctt	caccetteet	cacatcagct	tcctgctttt	cccagtgttt	1800
tactgtaaat	tgtgtagcat	atgacaaatc	ttgagctgac	tttcctcttc	acctgttatg	1860
gctggagtat	tttccagacc	tgaagggact	cacacttgtt	ttgatacttg	gatcacatct	1920
ccgtgaggtt	aggaaggtaa	atctaccaac	aggaagccct	gtactctgta	ttccaaggcc	1980
		cactgatcgg				2040
		cattgtctcc				2100
gcaacaggta	catacttgcc	ctttggaaat	ctaatacctc	tgggatacca	ttaagaggca	2160
ttttaattaa	acaaaagggc	ccttctaaat	gtgctattta	tttgacaata	actatcagat	2220
ttgccttaat	tttgtgttta	tagcatttat	caaaacgtat	cctcatagac	tttatgcaga	2280
ttaatatggt	caattgattt	ggataaaaga	aagtaatttc	agggtttgtt	tttaagccag	2340
gacaagaagt	gcaaatgcct	ctttgaagca	atttaggcta	aactgatttt	gaaatttcaa	2400
aatgttttat	tttactttgt	tttattaagc	caggacaaga	agtgcaaatg	cctctttgaa	2460
gcaattcagg	ctaggtaaac	cgattttgcc	atttcaaaac	gttttatttt	actttgwttt	2520
		ctgctgcaaa				2580
		gtacttcact				2640
		ttactaaaaa				2700
aatattatct	ggctccccaa	ttaaagtttg	attttattgt	cacaaaaaaa	aaaaaaaaa	2760
aaaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaa			2797 _.
<210> 75						
<211> 2703						
<212> DNA				-		
<213> Homo	sapiens					
<400> 75						
	ttcctacagg	tgaaacgcca	tcattaggat	tcactgtaac	gttagtgcta	60
					ctgcaacttg	120
gaaaaagagg	acctctcaga	aaactcacaa	tctagcatga	ttaagcatgt	cgcttggcta	180
					accattgatc	240
actocaatct	ctatcagece	cqaaataatq	aagtctgtta	ctctgatatt	ttttccattg	300
cctacttacc	tgaatccagt	cctgtatgtt	ttcttcaacc	caaagtttaa	agaagactgg	360
aagttactga	agcgacgtgt	taccaagaaa	agtggatcag	tttcagtttc	catcagtagc	420
						400

caaggtggtt gtctggaaca ggatttctac tacgactgtg gcatgtactc acatttgcag

ggcaacctga ctgtttgcga ctgctgcgaa tcgtttcttt taacaaagcc agtatcatgc

aaacacttga taaaatcaca cagctgtcct gcattggcag tggcttcttg ccaaagacct

gagggctact ggtccgactg tggcacacag tcggcccact ctgattatgc agatgaagaa

```
gattcctttg tctcagacag ttctgaccag gtgcaggcct gtggacgagc ctgcttctac
                                                                     720
cagagtagag gattcccttt ggtgcgctat gcttacaatc taccaagagt taaagactga
                                                                     780
actactgtgt gtgtaaccgt ttcccccgtc aaccaaaatc agtgtttata gagtgaaccc
                                                                     840
tattctcatc tttcatctgg gaagcacttc tgtaatcact gcctggtgtc acttagaaga
                                                                     900
aggagaggtg gcagtttatt tctcaaacca gtcattttca aagaacaggt gcctaaatta
                                                                     960
taaattggtg aaaaatgcaa tgtccaagca atgtatgatc tgtttgaaac aaatatatga
                                                                    1020
cttgaaaagg atcttaggtg tagtagagca atataatgtt agttttttct gatccataag
                                                                    1080
aagcaaattt atacctattt gtgtattaag cacaagataa agaacagctg ttaatatttt
                                                                    1140
ttaaaaatct attttaaaat gtgattttct ataactgaag aaaatatctt gctaatttta
                                                                    1200
cctaatgttt catccttaat ctcaggacaa cttactgcag ggccaaaaaa gggactgtcc
                                                                    1260
cagctagaac tgtgagagta tacataggca ttactttatt atgttttcac ttgccatcct
                                                                    1320
tgacataaga gaactataaa ttttgtttaa gcaatttata aatctaaaac ctgaagatgt
                                                                    1380
                                                                    1440
ttttaaaaca atattaacag ctgttaggtt aaaaaaatag ctggacattt gttttcagtc
attatacatt gctttggtcc aatcagtaat tttttcttaa gtgttttgtg attacactac
                                                                    1500
tagaaaaaaa gtaaaaggct aattgctgtg tgggtttagt cgatttggct aaactactaa
                                                                    1560
ctaatgtggg ggtttaatag tatctgaggg atttggtggc ttcatgtaat gttctcatta
                                                                    1620
atgaatactt cctaatatcg ttggctctac taatattttc caatttgctg ggatgtcacc
                                                                    1680
tagcaatagc ttggattata tagaaagtaa actgtggtca atacttgcat ttaattagac
                                                                    1740
gaaacgggga gtaattatga cacgaagtac ttatgtttat ttcttagtga gctggattat
                                                                    1800
cttgaacctg tgctattaaa tggaaatttc catacatctt ccccatacta ttttttataa
                                                                     1860
aagagcctat tcaatagctc agaggtigaa ctctggttaa acaagataat atgttattaa
                                                                    1920
taaaaataga agaagaaaga ataaagctta gtcctgtgtc tttaaaaaatt aaaaatttta
                                                                     1980
cttgattccc atctatgggc tttagaccta ttactgggtg gagtcttaaa gttataattg
                                                                    2040
ttcaatatgt tttttgaaca gtgtgctaaa tcaatagcaa acccactgcc atattagtta
                                                                     2100
ttctgaatat actaaaaaaa tccagctaga ttgcagttta ataattaaac tgtacatact
                                                                     2160
gtgcatataa tgaattttta tcttatgtaa attattttta gaacacaagt tgggaaatgt
                                                                     2220
ggcttctgtt catttcgttt aattaaagct acctcctaaa ctatagtggc tgccagtagc
                                                                     2280
agactgttaa attgtggttt atatactttt tgcattgtaa atagtctttg ttgtacattg
                                                                     2340
tcagtgtaat aaaaacagaa tctttgtata tcaaaatcat gtagtttgta taaaatgtgg
                                                                     2400
gaaggattta tttacagtgt gttgtaattt tgtaaggcca actatttaca agttttaaaa
                                                                     2460
attgctatca tgtatattta cacatctgat aaatattaaa tcataacttg gtaagaaact
                                                                     2520
cctaattaaa aggtttttc caaaattcag gttattgaaa atttttcatt ttattcattt
                                                                     2580
aaaaactaga ataacagata tataaaagtg ttaatctttg tgctatatgg tatgaaatac
                                                                     2640
                                                                     2700
2703
aaa
```

```
<210> 76
<211> 742
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (707)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (724)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (726)
<223> n equals a,t,g, or c
<400> 76
```

tctttgagaa ttggagggca	atagtgggtc gtttttcctt cttctttcgt caaccaaagg	tctccgcaac ggtagttctt	cagatgtaca ttattttat	tatttgaact taatctctgt	ctctttgtac atccttagat	60 120 180 240
cagtaataag	taagttgatt	atatacgage	tatqtaactt	atattttta	atggttggat	300
atcactgagt	tttttttt	aagaatttt	ttattgaggt	aaacttcaca	taacataaaa	360
ttaactattt	taaagtgaga	agttcagtgc	cacttagtat	tgttaacaat	gttgcataac	420
caccaccttt	atttaaagtt	ссававава	tattctcctc	taaaaggaaa	ccccatccca	480
ttaaggagat	actctccatt	ccttccttcc	tccagccccc	agcaaccacc	aatctgcttt	540°
ctatatatat	ggatttatct	attettacta	ttttatataa	atcgaattgt	atgagacctt	600
ttatatctaa	cttctttcac	ttagtagaag	tttttgagat	ttatttacat	agtagcatgt	660
atcaacactt	catttttatg	gccaaataaa	attotattat	gtgtttntag	cacaaaaaaa	720
	atgacceteg					742
aaananaaaa	acyacceccy	ag				•
<210> 77						
<211> 1825						
<212> DNA						•
<213> Homo	sapiens					
<400> 77						
ggcacgagca	tgtcacatgt	atacctatgt	aacacacctg	cacattctgc	acatgtatcc	60
cagaacttaa	attataataa	taaaaaaaga	ataattgggt	gatggcacat	ccaggtttgc	120
caaagacagt	cccagtttat	gctgttgtcc	tggcattatt	aataatgaca	ctgcctttaa	180
ctctcacaat	taatttggat	gataacttat	atggtaactc	tgctaaataa	aaaaaataaa	240
aattaccata	gtaacaggaa	cctacttgaa	atgatgcctc	tgtttctatt	ctggcttgaa	300
ttctgcattc	tttgaggatt	tgtagcctca	tgacagaatc	ctatctacag	gtgatgtatt	360
tcatatgatt	tttggctatt	tttttaacaa	tctcaagccc	aataatagcc	agtgatataa	420
ggaatgtagt	tactttctcc	ccactttctg	gcaagttaag	tttagccacc	tgattacaag	480
aagggacatt	cagaggtagg	atggcacaaa	gacacagggt	ccactggaga	tcactggaag	540
cagctgcagc	agggttaaga	gaagggagtc	ccagcgagtc	ttcagtcacc	acacactaac	600
atcatcagtg	aaaagttcct	gggcctgaag	atccagctat	gttgtttcta	gttgactatt	660
ttaagtgaca	gaacttggcc	caagcattga	ccattttggt	tcctcaataa	gcctgattca	720
accagggtca	cctttgaatc	tgtcctccac	ctttccdata	aacctatttt	atgcatcatt	780
cagtgagtta	tttatttatt	tacttttttg	ctgagaaaca	tgactagatt	taggaaaaat	840
gtagaatttt	acttttttt	caatattttc	tgggttttcc	agagttttca	cgtgtttcac	900
accttccttt	gcttcccacc	attccccttt	ctatttggaa	ctagagagac	atgagtttga	960
attctagctg	tgtaacctga	gtcagttatt	taacctcttt	ttgtttctgt	ttctttgtct	1020
gtaaatagca	aaaactacaa	ttaactttag	tccctgctgt	acaccaaatg	ttatcttgaa	1080
atattataca	tattatatgt	aattactact	gaaatgctct	aagatgccta	tgtgtgaatg	1140
gcattgttgt	aaagattaaa	taatataagg	gaagtgtctg	cttcagtgtc	tggcatataa	1200
taaaagctat	tatttttacg	attattttcc	atcttataga	agaattatcg	ttcttccctt	1260
ccaaagctaa	taaatggaca	tgtgtttatc	agacagaacg	taagagctgc	caaataaata	1320
gggaataggt	gctttcggga	gtctagggaa	ataaaggtca	gggaattgtt	cataaaattt	1380
agtacccata	aatagcctat	aagtagattc	cctagtttat	tctatgcagg	aaaataaagt	1440
tctacggagc	acagattcca	aaactaattg	gtcataaata	tcacctgaaa	gtttagaaaa	1500
tgtagcatca	tggacctctt	ttcataggtt	ctaaatctta	atatetgtgg	gatggtgcag	1560
gaatctagct	ttgctaagtg	ccctcagatg	actcttgctg	ctctaggcta	aaatacatgt	1620 1680
ggtttggctt	caatggacat	gttcctgaag	aatgtttgga	tgtcacacat	tcatatttag	1740
tatgagagat	gaggtcctcc	tctcatcatt	ttcttaggtt	CTCTCTCTC	cactccttac	1800
			aaattagcta	tacatttgtt	tcattataaa	1825
aaagaaagaa	gataaaaaaa	aaaaa				1023
						•

<210> 78

<211> 1674

<212> DNA

<213> Homo sapiens

```
<400> 78
ggccacgaga gtatctgcgg cagctgcagg tcctggattt atttctcgat tcgctgtcgg
                                                                     60
aggagaatga gaccctggtg gagtttgcta ttggaggcct gtgcaacctg tgcccagaca
                                                                     120
gggccaacaa ggagcacatc ctgcacgcag gaggtgtccc actcatcatc aactgcctat
                                                                     180
ccagccccaa tgaggagacg gtgctgtctg ccatcaccac gctcatgcac ctgagcccgc
                                                                     240
                                                                     300
cgggccgcag ctttctccca gagctgaccg ccacgcccgt ggtgcagtgc atgcttcgct
                                                                     360
tetecetete ggecagegee aggtteegga acetggeaca gatetteetg gaggaettet
gctcccccg ccaagtggcc gaggcccgca gccggcaggc gcaattttgc cctgggtatc
                                                                     420
ccactgccga ggagcgtggc cccacggcag cgctgatcca tggagactgc gagaccgtgg
                                                                     480
                                                                     540
cacccctact gctggggacc acagtcctga tgtggacgca gggaacgggg agcacatact
                                                                     600
gcccattgg tgccttttca gccatctgaa aggcgggttc tttcagcagg acaggcattt
                                                                     660
acactgatga aacgccactg ggagtgagga agccagactc cagagacacg gagaagatca
                                                                     720
aactggagct gcgttcatag gctggcactc tcaatcctac atcaggtgcc accaccacca
                                                                     780
gactcaggcc ctggtgtaag aagcggccaa gtgcctggac ccagaggctt tgcaggacag
                                                                     840
tgttctcagg agctgggcct gaggcttagg agagctgcct tcgctgcagg aaatcaggga
ttatccctta acagaagtgt ctggagtagt tttcaggtat aggaatgaga tgcctcgtgg
                                                                     900
                                                                     960
tgaaaggatc tcaccctggg aagatgtggt gccccctcca gggctctgga ggatggatgc
ctccccagg ggctctccaa gctgggcatt tgggcctggt ggatgccaac ctggataacc
                                                                    1020
tgtggcccag cattgactgt ccacccagcc ttgctgttag gcaccatgac tccaaagatg
                                                                    1080
aagatgtggt ccctgccctt gagtgacagc cccagggact taatgtggcc atcgggcatc
                                                                    1140
                                                                    1200
aagcacaagg ccatgcaggt gatgatacgt cggaatagag gcaccagccc tggtaactgc
atcttctccc cttgccaccc catggccccg gctgaaagct tcggccctcc tctgctgtca
                                                                    1260
                                                                    1320
ctcaatgatg gggagcccta ccccagaagt gtatcccacg agggcatcag ggacgcagtg
                                                                    1380
agtgttgctc aagggagtca ggaagagacg gcaacgtaaa ggatgtggct ccatgtccat
                                                                    1440
ggtgcccct ggtcaacata aggagcgtgg gatccgatgg aaaggtggag ctcagggaaa.
atgggggtcc ttgcctctcg tgtaccccct caaggctgac cccttagatg gcccaggaat
                                                                    1500
ggcaggtgct acaaaaatgg tacccacgtg ggcatggaaa tggggcagat taggggacca
                                                                    1560
ctggactcag aggggaggga agggctcatc agcacccgct cagggagcct gtccctttat
                                                                    1620
                                                                    1674
<210> 79
<211> 2191
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1327)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1334)
<223> n equals a,t,g, or c
<400> 79
ccttctctaa aaaagcaaac aggcaaaact tcatgagaat cttgatcatg ttaaaatttt
                                                                      60
120
tcaacatttc ctccacccat atcatcactc cttagcatct ttattccatc aaaactttct
                                                                     180
                                                                     240
accepttgae attetetgtg cagttttgaa aattaceete teageattet etgtteacee
                                                                     300
ccacacctag accctgacct ctagtcaatt ctactaccca ggggtgtcca cggttccagc
                                                                     360
ctcctccatg aagcccagtt ctatgggctc actcctctgg gtaagtggga gcccccagct
atcatcctca ttgtatagaa aaccaactct gtgatgctac ctgcccctct tccccttctc
                                                                     420
                                                                     480
tcctgaaaga gggctggggt agaggtggga ggactggtta tggccctggc cgggtctgta
                                                                     540
ttcgtactgg gaggagtatt ggtactctgt gtagaaagaa atggggaggg ggaaatgggg
                                                                     600
tggcctcagc atctccctaa gtcccagcct ttaagtcctc ctgttgcagt tcgtcgctgc
```

```
660
agcttcgaga ggagttggat cgatcttctt gtggaaacgt cctcttcaat ggttacctgc
                                                                      720
cgccaacagg taggcactcc caatggaatg gaggggcgcg gaggtgggcc aaagactaca
                                                                      780
840
gtgtagttcc ctatcagatg cttgggctga tgcttggaaa ggaagttgga cacagcattt
cccatgaaac aatgggccaa ctaactcttg aagctcaaaa agatgtcctt ggaaccccat
                                                                      900
                                                                       960
ggggaatttg ttatcccggg tttgggtttc ttttgttagg gggggctttg ggaaaaactg
                                                                      1020
gggattcctc cgtatggaag gggaaaaaat attaaatagg aagttattga cattaatgcc
                                                                      1080
catgatagcc accccactgg gccatggaag gtatgcccca gtgggtattg gaactaggct
                                                                      1140
tttctgattg gtagaagtaa cagagtaggg aaatttcatc tacagcttta tttccctaac
                                                                      1200
tgcagtcagc acctgtacct tcatgaaagt tgccagatat aaagatctgt agtagtactt
                                                                      1260
ttccaactta gttttatcct gttttcccga aaaacaatca tttatttatt tatttattta
tttaatttta tgagacaggg tctggctttg tcacccaggc tggagtgcag tggtgcgatc
                                                                      1320
                                                                      1380
ttggctncac tgcnacctct gcctctcaga ttcaagccat ccttccacct cagctctgcc
                                                                      1440
actgagtage tgagactaca ageactegee accatgeeeg getaattaaa aaaataataa
                                                                      1500
tcattttaaa tgcaagcttt atattataaa tacaaagtaa acatgaaaat aaaacccaaa
catagoagtg ttattaaact ctggcctgta gcagtggctc acacctgtaa tcctagcagt
                                                                      1560
ttggaggccg agacaggtgg attacttgag acctggagtt tgagaccagc ccaggtgaca
                                                                      1620
cagcaagacc tcatctctac taaaaataaa aaaaaattag ccaggtgtgg tggtatgcac
                                                                      1680
ctgtggtccc agctacttag gatgctggag tgcgaggatc gcttgagccc aggaggtcaa
                                                                      1740
ggctgcagtg aactatgatc actcattaca ccccagcctg ggtgacagag cgagatgctg
                                                                      1800
tctcaaaaca aaacaaacg aaaaacaact ctggctagat gctattgctt gccaagggtg
                                                                      1860
                                                                      1920
cagtetteca tttattaaaa gtgaaaatta gggeeaggea cattggetea tgeetgtaat
                                                                      1980
cccagcactt tgggaggctg aggtgggtgg atcacctgag gtcaggagtt cgagaccagc
ctggccaaca tggtgaaacc ttatctctgc caaaaatata aaagattagc catgtgtcgt
                                                                      2040
                                                                      2100
ggtgggtgct tgtaatctca gctacttggg aggctgaggc aggagaatca cttgaaccca
ggaggcagag gttgcagtga gccaagattg tgccattgca ctccagcctg tgcaacgagc
                                                                      2160
                                                                      2191
gaaactccaa ctcaaaaaaa aaaaaaaaaa a
<210> 80
<211> 1335
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1287)
<223> n equals a,t,g, or c
<400> 80
                                                                        60
ggatatatcc agggctgcgg attttccccc cttcaggttt aaatgttcct gtttttctac
ctttccctcg cagtatacgc tcaacggcaa gawagtggaa gttgccgtca aacagatcat
                                                                       120
cgctggaaaa gccgtggagc aaggaggtgc tttctcgaac cccgagaccc tggatctgta
                                                                       180
                                                                       240
ccgggacatc cctgagctgc agggcttctg agtcagactg gctggcgtgt cactcagccg
cacccgtgtg cactgtaact tttgtgtgct caagaaatta tacagaaacc tacagctgtt
                                                                       300
gtaaaaggat gctcgcacca agtgttctgt aggcttgggg agggatcgtt tctctgtttt
                                                                       360
                                                                       420
gttaaatctg gtgggtacct ggatcttcca cacgagtggg attctggcct tcagagacca
                                                                       480
ggagggagtg tctgggccgc agtgtggcac tgtggtgaga gtgtgtgtct ttgcacacac
                                                                       540
agtgcagcgg gaacggtggg gctggctggt gctgaagaca gacacactcc tgagccaagg
                                                                       600
tettgtette aaceteceeg teeegttgte ceattttget etgtgaaggt geaaateeet
ttcttccctt cccatctcag gctctcctgt tttccctcag ggtccagtat gcctttgagc
                                                                       660
                                                                       720
tttagctgtt agaaaggaac ccccgtgact tgacacagct ttcacagctg gctgctagga
                                                                       780
ccggcgggct gggtgttcac gtgtgtctgt gtcatggatg caatgcaggc cctggaggac
                                                                       840
tgtgcgtcac ccgtcaacca gagcgtgcct ccgggccagc ttccctccaa ggaatgagtg
                                                                       900
gatttcatac aggatctctt tattgcacag actgaatggc tttacatgtt tctaatgtga
attaggcatg tgaagcagtg ggtgtccacc cgtgtccctc atgggtgagc cctccagctg
                                                                       960
tgagcccagg cagtgtggtc accgagtgag gaccctcctc accaggaacc gcatccctgt
                                                                      1020
```

gctgcctcca cctgagagtt gctagggggt tcttgtcgag atcatgtcat cagcacccct

aagtcaagtc acgggtttcc atagccaggc agttggtatg tacaattcag ttcagcgt gaacttgtat ctctaatctg atgtccattt ttatatttt tgaaactgag cacaatga tcctttcttg aatcattttc cttttggatt ataaaaatat gggggaaagt gctatgat attttatgca ataaatgtat acatgtntgc acatgcaccc atgctgaaaa aaaaaaaaaa aaaaa	ga 1260
<210> 81 <211> 1867 <212> DNA <213> Homo sapiens	
<400> 81	
cccacgcgtc cgggccacag cagagacagt ggagggcagt ggagaggacc gcgctgtc	cct 60
gctgtcacca agagctggag acaccatctc ccaccgagag tcatggcccc attggccc	tg 120
cacctcctcg tectegtece catcetecte ageetggtgg ceteceagga etggaagg	gct 180
gaacgcagcc aagacccctt cgagaaatgc atgcaggatc ctgactatga gcagctgc	etc 240
aaggtggtga cctgggggct caatcggacc ctgaagcccc agagggtgat tgtggttg	ggc 300 acc 360
gctggtgtgg ccgggctggt ggccgccaag gtgctcagcg atgctggaca caaggtca	
atcctggagg cagataacag gatcgggggc cgcatcttca cctaccggga ccagaaca	- 3
ggctggattg gggagctggg agccatgcgc atgcccagct ctcacaggat cctccaca	
ctctgccagg gcctggggct caacctgacc aagttcaccc agtacgacaa gaacacg	
acggaggtgc acgaagtgaa gctgcgcaac tatgtggtgg agaaggtgcc cgagaag	-
ggctacgcct tgcgtcccca ggaaaagggc cactcgcccg aagacatcta ccagatggctcaaccagg ccctcaaga cctcaaggca ctgggctgca gaaaggcgat gaagaag	•
gaaaggcaca cgctcttgga atatcttctc ggggagggga	ata 780
cagettetgg gagacgtgat gteegaggat ggettettet ateteagett egeegag	, ,
ctccgggccc acagctgcct cagcgacaga ctccagtaca gccgcatcgt gggtggc	-
gacetgetge egegegeget getgageteg etgteeggge ttgtgetgtt gaaegege	ecc 960
gtggtggcga tgacccaggg accgcacgat gtgcacgtgc agatcgagac ctctccc	ccg 1020
gcgcggaatc tgaaggtgct gaaggccgac gtggtgctgc tgacggcgag cggaccg	gcg 1080
gtgaagcgca tcaccttctc gccgccgctg ccccgccaca tgcaggaggc gctgcgg	agg 1140
ctgcactacg tgccggccac caaggtgttc ctaagcttcc gcaggccctt ctggcgc	gag 1200
gagcacattg aaggcggcca ctcaaacacc gatcgcccgt cgcgcatgat tttctac	ccg 1260
ccgccgcgcg agggcgcgct gctgctggcc tcgtacacgt ggtcggacgc ggcggca	gcg 1320
ttcgccggct tgagccggga agaggcgttg cgcttggcgc tcgacgacgt ggcggca	ttg 1380
cacgggcctg tcgtgcgcca gctctgggac ggcaccggcg tcgtcaagcg ttgggcg	gag 1440
gaccagcaca gccagggtgg ctttgtggta cagccgccgg cgctctggca aaccgaa	aag 1500
gatgactgga cggtccctta tggccgcatc tactttgccg gcgagcacac cgcctac	ccg 1560
cacggctggg tggagacggc ggtcaagttg ctgcgcgccg ccatcaagat caacagc	cgg 1620
aaggggcctg catcggacac ggccagcccc gaggggcacg catctgacat ggagggg	cag 1680
gggcatgtgc atggggtggc cagcagcccc tcgcatgacc tggcaaagga agaaggc	agc 1740
caccetecag tecaaggeea gttatetete caaaacaega eccaeaegag gaceteg	cat 1800
taaagtattt tcggaaaaaa aaaaaaaaaa aaaaaaaaa aaaaaaaaa	aaa 1860
aaaaaaa	1867
	•
<210> 82	
<211> 984	
<212> DNA	
<213> Homo sapiens	
<400> 82	
gaatteggea egageecage ggaageeaag ceaceaggee eeceagegte caegeg	gagc 60
atgaacattg aggatggcgc gtgcccgcgg ctccccgtgc cccccgctgc cgcccg	gtag 120
gatgteetgg ecceaegggg cattgetett cetetggete tteteceeae eeetgg	gggc 180
cggtggaggt ggagtggccg tgacgtctgc cgccggaggg ggctccccgc cggcca	cctc 240
ctgccccgtg gcctgctcct gcagcaacca ggccagccgg gtgatctgca cacgga	gaga 300

mctggccgag	gtcccagcca	gcatcccggt	caacacgcgg	tacctgaacc	tgcaagagaa	360
coocatccao	gtgatccgga	cggacacgtt	caagcacctg	cggcacctgg	agattctgca	420
gctgagcaag	aacctggtgc	gcaagatcga	ggtgggcgcc	ttcaacgggc	tgcccagcct	480
caacacacta	gagetttttg	acaaccggct	gaccacggtg	cccacgcagg	ccttcgagta	540
cctgtccaag	ctgcgggagc	tetagetgeg	gaacaacccc	atcgagagca	tcccctccta	600
caccttcaac	cacataccct	cactacaaca	cctggacctg	ggcgagctca	agcggctgga	660
atacatctcg	gagggggcct	tcgargggct	ggtcaacctg	cgctacctca	acctgggcat	720
gtgcaacctc	aaggacatcc	ccaactgacg	gccctggtgc	gcctggagga	gctggagctg	780
tcgggcaacc	ggctggacct	gatccgcccg	ggctccttcc	agggtctcac	cagcctgcgc	840
aagctgtggc	tcatgcacgc	ccaggtagcc	accatcgagc	gcaacgcctt	cgacgacctc	900
aagtcgctgg	aggageteaa	cctgtcccac	aacaacctga	tgtcgctgcc	ccacgacctc	960
	tgcaccgcct					984
•	-					

<210> 83 <211> 2664 <212> DNA

<213> Homo sapiens

. <400> 83

60 ggttgctggc ccaggtgagc gggcgcgctg gtccaggtga gcgggcgcgt ccccgcgacg gcgctgcctg cccgaggcgg ttcacgtaaa gacagcgaga tcctgagggc cagccgggaa 120 ggaggcgtgg atatggagct ggctgctgcc aagtccgggg cccgcgccgc tgcctagcgc 180 gtcctgggga ctctgtgggg acgcgcccg cgccgcggct cggggacccg tagagcccgg 240 cgctgcgcgc atggccctgc tctcgcgccc cgcgctcacc ctcctgctcc tcctcatggc 300 360 cgctgttgtc aggtgccagg agcaggccca gaccaccgac tggagagcca ccctgaagac 420 cateeggaae ggegtteata agatagaeae gtaeetgaae geegeettgg aceteetggg aggcgaggac ggtctctgcc agtataaatg catgacggat ctaagccttt cccacgttat 480 ggttataaac cctccccacc gaatggatgt ggctctccac tgtttggtgt tcatcttaac 540 attggtatcc cttccctgac aaagtgttgc aaccaacacg acaggtgcta tgaracctgt 600 ggcaaaagca agaatgactg tgatgaagaa ttccagtatt gcctctccaa gatctgccga 660 gatgtacaga aaacactagg actaactcag catgttcagg catgtgaaac aacagtggag 720 780 ctcttgtttg acagtgttat acatttaggt tgtaaaccat atctggacag ccaacgagcc 840 gcatgcaggt gtcattatga agaaaaaact gatctttaaa ggagatgccg acagctagtg acagatgaag atggaagaac ataacctttg acaaataact aatgttttta caacataaaa 900 ctgtcttatt tttgtgaaag gattattttg agaccttaaa ataatttata tcttgatgtt 960 aaaacctcaa agcaaaaaa gtgagggaga tagtgagggg agggcacgct tgtcttctca 1020 1080 ggtatcttcc ccagcattgc tcccttactt agtatgccaa atgtcttgac caatatcaaa aacaagtgct tgtttagcgg agaattttga aaagaggaat atataactca attttcacaa 1140 ccacatttac caaaaaaaga gatcaaatat aaaattcatc ataatgtctg ttcaacatta 1200 1260 tcttatttgg aaaatgggga aattatcact tacaagtatt tgtttactat gaaattttaa 1320 atacacattt atgcctagaa ggaacggact ttttttttct attttaatta cacataatat 1380 gtaattaaag tacaacataa tatgttgttt ctctgtagcc cgttgagcat atgagtaagt cacatttcta ttaggactac ttmcaaggac aaggtttcca tttttccagt tgtaaaattg 1440 1500 gaaccatcag ctgataacct cgtagggagc aaccccagga tagctaagtg ttatgtaata tgcctagaag gtgatgtgaa tgcgattcag aagcatagcc actcccattt tatgagctac 1560 tcacatgaca aatgtcatct tttgctataa cctttgccaa gttagagaaa agatggattt 1620 1680 aatgagataa atgaaaagat atttamccta atatatcaag gcactatttg ctgttatgct 1740 ttgttattta tttcccagca cttgttcctt attgtagatt ttttaaagac tgtaaccttt 1800 tactaactgt ggtcttacta aaatttgtgc ttgatactgc ttttcaaaaa gcctttaatt agagccaaaa ggatggaaaa ggcaagatat aaatgccttt tatagatctc ttatttacat 1860 tgaaaattat taccatatgt ttagagcaaa tccaagaaaa cttcaacagc ttctgaagat 1920 gtctatgaat gttgaaaact tttcaatctc ttggaatgct cagttatgtt cctagaccgg 1980 2040 tctttgctga ctactggttg ttaacctttc cctagcctgg gacctcaagc catatatatc 2100 ctttgggtga cccatggcca aagttattaa gatgaactga ctttcaaagt cagagaagga 2160 cagcataggg agaggcggtt atttgtaagt cattacaggt agaacagggc agaaggaaaa gtatgttctg gagaaagggc catgttccta actttggaga tatgtcattg ccgggaacct 2220 agtatcttcc aacttgaatt ggtggcagct gttccagtga gacaaggcac atgtatgcct 2280

			52			
tatttcaaac cggctgggca tcacttgagg aaaatacaaa ctgaggtacg	ctttataaaa cagtggctca tcaggagttt aatttgccgg	gggtttccac ggtacagttt cacctataat gagaccagct gcgtggtggc gaacccagga cgag	tgtaagccat cccagcactt ggccaacatg gcatgcttat	tattaataat gggaggctga gtgaaacagc agtctcagct	taatgettat ggeggttgga gtetetaeta acteaggaag	2340 2400 2460 2520 2580 2640 2664
<210> 84 <211> 1328 <212> DNA <213> Homo s	sapiens			·		
ttaaagttcc gtgtgacgag actgtgtact agccgcaaac cactccctc gacagtggga agctctgcct cgggagatct ggaggacagc ctgtgctcct tggacactcc gatccattt tcccaattgt ccaacttact ctggaattgc atagttcca acgagctgcc taagccagac cgggcagcta atcacaccag	gggaatcaaa agaacgagat gcatgaatga aggaaaaggc cctgctccgt aacaggcact gtagcaggag ttttcagcta tgagcttcca acctttgtgc ggactcctt ctgtgttagt ggaaccaaag tacagtttag tcatgagtct cacagtctc aggcctcct gctataaagc agtacatgtc agtacatgtc	gaggtccgca gatcaactcc ttaccttcct ggctgagata ctgcgagcag gtcgcctcac grctccaggc ccctgaggag gggcataaac ctggtgctgc cttcttgggc gctgccgttc gtatgtactg cagtgaccag agacagtact tccgcatgat atcaatgagc tgccctgtct ccgcaagctg aagagggaca tctgcctctg ggtttccaa	cactgaggac gaattaaaaa gttgatgttg ccggccctgg acaagttctg ctcagccctt gaggaggaag tgtgcactga tgggccgcc ctggcagatt tgagtgtgcg tgccctccc gttccaggga ttgcaaagaa ctctcctgaa tcccacctct gggaggttga ctaccttgct ggagcccaga tttcagtgt	aaatggacct wcwgactccc ctctgggaat cgggggctga ggagcagcag cccagaggcc aggatgtgct actgtctgcc gcctgtggga cacctcaggc gaaggcagta tggctcttga aaatgatgtc aaggatgacc cgaggaagcc ccagccagcc ggccacagtg ttcacctgta agagacactg ggcttttggac	gtaattccgg tgcgacaagg cctgattgag taacccagar tgaggaagag ggggggttcc gaaatacgtc gagagcagct atggggctct cagaagcccc ctcaaaatga ggctctggag atgtggtgct gccaggtgca tgttcaaaa tagaaagcaa tatagactgg ccttggtccc aggacaagag aggaatatat	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1328
<210> 85 <211> 1342 <212> DNA <213> Homo s	sapiens					
aattgatctg ggtcatgcca ttgggaagag gatggacttt tatctgcgtc attttcctgt tttccatacg tactgagtac accttcttct gaaccaatcc	tctctgcagc ggcacacact gaatgccaga ctggtcctct tgctcgaaaa ataattccag agaaaccaca acctggggaa gccctatctg tggcattata	gctgcccaag gctgccggct tcttgttcta cccatagctt aatgtcttca ccttcattgt gtatttggct ctgctaggtg acaaaagcaa	ctcctggaag aggagctgct gaaaattacc cctggcttcg gaaaggcctg gagagccrtg cctgcacctg actgtcagga taaacctgtt atgaattatt	aaccatgtcc gtttgaatta caaccaagag gtgctgatgg gccaggggag catggattgc gtcttgcaag gctggagttg tttttcacc	ggcagctact tctgtgaatg aaatctgcag gtcttgttct gagcacagat ttcattacct ggatggttta	60 120 180 240 300 360 420 480 540 600 660 720

<221> SITE

```
ctcgatccaa gcactgcagt gtgtgtaact ggtgtgtgca ccgtttcgac catcactgtg
                                                                         780
tttgggtgaa caactgcatc ggggcctgga acatcaggta cttcctcatc tacgtcttga
                                                                         840
cettgacgge eteggetgee accgtegeea ttgtgageae caettttetg gtecaettgg
                                                                         900
tggtgatgtc agatttatac caggagactt acatcgatga ccttggacac ctccatgtta
                                                                         960
 tggacacggt ctttcttatt cagtacctgt tcctgacttt tccacggatt gtcttcatgc
                                                                        1020
tgggctttgt cgtggttctg agcttcctcc tgggtggcta cctgttgttt gtcctgtatc
                                                                        1080
tggcggccac caaccagact actaacgagt ggtacagagg tgactgggcc tggtgccagc
                                                                        1140
gttgtcccct tgtggcctgg cctccgtcag cagagcccca agtccaccgg aacattcact
                                                                        1200
cccatgggct tcggagcaac cttcaagaga tctttctacc tgcctttcca tgtcatgaga
                                                                        1260
ggaagaaaca agaatgacaa gtgtatgact gcctttgagc tgtagttccc gtttatttac
                                                                        1320
                                                                        1342
acatgtggat cctcgttttc ca
<210> 86
<211> 1154
<212> DNA
<213> Homo sapiens
<400> 86
                                                                          60
 aagacaggaa aagctccagg ccgtggttct caaagtgtgg tccctggaca gcagcaacat
                                                                         120
 cacctaggag cctgttaggg aaggcacagc ctcaggccct gccccagacc tgcagaatca
 gaaactctgg ggtgaggcct ggttatctgc tgtaacagac cttccagtgg gttctgatgc
                                                                         180
                                                                         240
 cctctagagc aggagaacca ctagcttaga ggttgcagta tgtttggcat cttgccattt
                                                                         300
 gtgttagttc agaggaatgg ctgaccccca tgtctcattt ctaagcttca ggcagctttt
 ctcctgggca gctgtcattc tgttgagggg aatcctgggg actgtggctc ctcctccctg
                                                                         360
                                                                         420
 tccgtgtgtc cttgatctgg cagtctaccc ccttcatctc cccgtggagg ctccatgcct
                                                                         480
 agaggtggtc ttcaaacaga agaatggcaa agataattgt ctcgtgtttt accctgaccc
                                                                         540
 cattccttta agagggtcac ttcttggccc attcatttaa aaaccaatgt catagttctg
                                                                         600
 tgattccacc tatcagacag tgccacgtcc aaaggcgggg ctctyacctc cctggraaga
                                                                         660
 gagactgttg ctgtctgtgc ttcctgtgtt ctccagtccc acgctcccac ggacccacgc
 ccttggagac tccctcrgtg tcccagggct tctggtgtgt tcagagacct ccacactcaa
                                                                         720
                                                                         780
 cgaccactgg tgctgcagaa gggccggtgc ttacattcca attaacagac gcttttccca
 tctaatgcct cttgccttct cctaacacca cctcgggagt gtttatgtct attctaagtg
                                                                         840
 aatttcactg tgtgaaaaaa ttcacacctg ttgtcccagc gatttgggag gccggggcgg
                                                                         900
                                                                         960
 gtgtatcatt tgagcccagg agtttgaggc tagcctgggc aggatggtga aaccccgtct
                                                                        1020
 ctataaagaa attttaaaaa ttagctgggc atagtggcac gtgcctgtag ttccatctac
                                                                        1080
 tggggaggct ggggtgggag gatcgcatga gcccgggagt ttgaggctgc agtgagctgt
                                                                        1140
 gatcgcagca ctgcactcca gtctgggcaa cagagcaaga ccctgtctct taaaaaaaaa
                                                                        1154
 aaaaaaact cgag
<210> 87
<211> 1197
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (573)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1177)
<223> n equals a,t,g, or c
<220>
```

```
<222> (1185)
<223> n equals a,t,g, or c
<400> 87
 aagacaggaa aagctccagg ccgtggttct caaagtgtgg tccctggaca gcagcaacat
                                                                          60
 cacctaggag cctgttaggg aaggcacagc ctcaggccct gccccagacc tgcagaatca
                                                                         120
 gaaactctgg ggtgaggcct ggttatctgc tgtaacagac cttccagtgg gttctgatgc
                                                                         180
 cctctagagc aggagaacca ctagcttaga ggttgcagta tgtttggcat cttgccattt
                                                                         240
 gtgttagttc agaggaatgg ctgaccccca tgtctcattt ctaagcttca ggcagctttt
                                                                         300
 ctcctgggca gctgtcattc tgttgagggg aatcctgggg actgtggctc ctcctcctg
                                                                         360
 tccgtgtgtc cttgatctgg cagtctaccc ccttcatctc cccgtggagg ctccatgccw
                                                                         420
 agaggtggtc ttcaaacaga agaatggcaa arataattgt ctcgtgtttt accctgaccc
                                                                         480
 cattccttta agagggtcac ttcttggccc attcatttaa aaaccaatgt catagttctg
                                                                         540
 tgattccacc tatcagacag tgccacgtcc aangcggggc tctcacctcc ctgggaagag
                                                                         600
 agactgttgc tgtctgtgct tcctgtgttc tccagtccca cgctcccacg gacccacgcc
                                                                         660
                                                                         720
 cttggagact ccctcagtgt cccagggctt ctggtgtgtt cagagacctc cacactcaac
                                                                         780
 gaccactggt gctgcagaag ggccggtgct tacattccaa ttaacagacg cttttcccat
                                                                         840
 ctaatgcctc ttgccttctc ctaacaccac ctcgggagtg tttatgtcta ttctaagtga
 atttcactgt gtgaaaaaat tcacacctgt tatcccagca atttgggagg ccgaggcggg
                                                                         900
                                                                         960
 tgtatcattt gggcccagga gtttgagact agcctgggca agatggtgaa accccgtctc
 tataaagaaa ttttaaaaat tggctgggca tagtggcgcg tgcctgtagt tccatctgct
                                                                        1020
 ggggaggctg gggtgggagg atcgcatgag cccgggagtt tgaggctgca gtgagctgtg
                                                                        1080
 atcgcggcac tgcactccag tctgggcaac agagcaaaac cctgtctctt aaaaaaaaa
                                                                        1140
 aaaactcgag ggggggcccg gtacccaatt cgccctnats agtgnagtcg tattaca
                                                                        1197
<210> 88
<211> 910
<212> DNA
<213> Homo sapiens
<400> 88
                                                                          60
 ggcagagctg gccttcgact cgctatgtcc actaacaata tgtcggaccc acggaggccg
 aacaaagtgc tgaggtgagg accccagcgt cgtgggcacg ggttcgggtt gtgggtgtgg
                                                                         120
 ateggggeec tgggaagege ctgtctatec egggggeagg acetgagege ecetgaeect
                                                                         180
 cgagcctgtc gcaggtacaa gccccgccg agcgaatgta acccggcctt ggacgacccg
                                                                         240
                                                                         300
 acgccggact acatgaacct gctgggcatg atcttcagca tgtgcggcct catgcttaag
                                                                         360
 ctgaagtggt gtgcttgggt cgctgtctac tgctccttca tcagctttgc caactctcgg
                                                                         420
 agctcggagg acacgaagca aatgatgagt agcttcatgt gagacttgcc ctacagaaca
                                                                         480
 agtgactctt gagtaagggg tggggggacc ccagcctggc catcctagac tgacacctct
                                                                         540
 ctcctgtctt catgctgtcc atctctgccg tggtgatgtc ctatctgcag aatcctcagc
                                                                         600
 ccatgacgcc cccatggtga taccagccta gaagggtcac attttggacc ctgtctatcc
 actaggeetg ggetttgget getaaacetg etgeetteag etgeeateet ggaetteeet
                                                                         660
                                                                         720
 gaatgaggcc gtctcggtgc ccccagctgg atagagggaa cctggccctt tcctagggaa
                                                                         780
 caccctagge ttacccctcc tgcctccctt cccctgcctg ctgctggggg agatgctgtc
                                                                         840
 catgtttcta ggggtattca tttgctttct cgttgaaacc tgttgttaat aaagtttttc
                                                                         900
 actctgaaaa aaaaaaaaa aaaaaaaaac tygrgggggg gcccggaacc caattcsccg
                                                                         910
 gatagtgagt
<210> 89
<211> 1076
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1029)
```

```
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1037)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1040)
<223> n equals a,t,g, or c
<400> 89
 ggcacgaggg gaaagccatg ctcccaggac tccttccttg cagccttaaa tcggtctgta
                                                                          60
                                                                         120
 cggaaaattc cgcgccttag aaacccacgc ttgggtgtaa cttattattg ttcttcctga
                                                                         180
 cctacttcct gtttatcact tccgggttca tcattttggc atttcggtga tcgggttgga
 actattgaag cccgctttca ggttcttttc cccattttcc ctttgaaagg aagacttctg
                                                                         240
 gcttctccta aatctccgtt ctctgggtaa ggggagtcca agcctctgtc atgaggaacg
                                                                         300
 gaaatgcgag ggcctcgggt gttactctaa aatccgccct cagcttgcac gccggaagct
                                                                         360
 gcgattcctg cagcggaaga ggcgtgatct ggccttcgac tcgctatgtc cactaacaat
                                                                         420
                                                                         480
 atgtcggacc cacggaggcc gaacaaagtg ctgaggtaca agcccccgcc gagcgaatgt
                                                                         540
 aacceggeet tggacgacce gacgeeggac tacatgaacc tgetgggeat gatetteage
                                                                         600
 atgtgcggcc tcatgcttaa gctgaagtgg tgtgcttggg tcgctgtcta ctgctccttc
 atcagetttg ccaacteteg gageteggag gacacgaage aaatgatgag tagetteatg
                                                                         660
 ctgtccatct ctgccgtggt gatgtcctat ctgcagaatc ctcagcccat gacgccccca
                                                                         720
                                                                         780
 tggtgatacc agcctagaag ggtcacattt tggaccctgt ctatccacta ggcctgggct
 ttggctgcta aacctgctgc cttcagctgc catcctggac ttccctgaat gaggccgtct
                                                                         840
 cggtgcccc agctggatag agggaacctg gccctttcct agggaacacc ctaggcttac
                                                                         900
 ccctcctgcc tcccttcccc tgcctgctgc tgggggagat gctgtccatg tttctagggg
                                                                         960
                                                                        1020
 tattcatttg ctttctcgtt gaaacctgtt gttaataaag tttttcactc tgaaaaaaaa
                                                                        1076
 aaaaaaaana raaaacncgn gggggggccc ggaacccaat tcsccggata gtgagt
<210> 90
<211> 1842
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (67)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (98)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (212)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1838)
<223> n equals a,t,g, or c
```

```
<400> 90
gegacegege cetteageta getegetege tegetetget teeetgetge eggetgegea
                                                                        60
tggcttnggc gttggcggcg ctggcggcgg ctcgagcngc ctgcgsagcc ggtaccagca
                                                                       120
gttgcagaat gaagaagagt ctggagaacc tgaacaggct gcaggtgatg ctcctccacc
                                                                       180
ttacagcagc atttctgcag agagcgcaca tnattttgac tacaaggatg agtctgggtt
                                                                       240
                                                                       300
tccaaagccc ccatcttaca atgtagctac aacactgccc agttatgatg aagcggagag
                                                                       360
gaccaaggct gaagctacta tccctttggt tcctgggaga gatgaggatt ttgtgggtcg
ggatgatttt gatgatgctg accagctgag gataggaaat gatgggattt tcatgttaac
                                                                       420
tttttcatg gcattcctct ttaactggat tgggtttttc ctgtcttttt gcctgaccac
                                                                       480
                                                                       540
 ttcagctgca ggaaggtatg gggccatttc aggatttggt ctctctctaa ttaaatggat
                                                                       600
cctgattgtc aggttttcca cctatttccc tggatatttt gatggtcagt actggctctg
 gtgggtgttc cttgttttag gctttctcct gtttctcaga ggatttatca attatgcaaa
                                                                       660
 agttcggaag atgccagaaa ctttctcaaa tctccccagg accagagttc tctttattta
                                                                       720
                                                                       780
 ttaaagatgt tttctggcaa aggccttcct gcatttatga attctctctc aagaagcaag
                                                                       840
 agaacacctg caggaagtga atcaagatgc agaacacaga ggaataatca cctgctttaa
 aaaaataaag tactgrtgaa aagatcattt ctctctattt gttcctaggt gtaaaatttt
                                                                       900
 aatagttaat gcagaattct gtaatcattg aatcattagt ggttaatgtt tgaaaaagct
                                                                       960
 cttgcaatca agtctgtgat gtattaataa tgccttatat attgtttgta gtcattttaa
                                                                      1020
 gtagcatgag ccatgtccct gtagtcggta gggggcagtc ttgctttatt catcctccat
                                                                      1080
 ctcaaaatga acttggaatt aaatattgta agatatgtat aatgctggcc attttaaagg
                                                                      1140
 ggttttctca aaagttaaac ttttgttatg actgtgtttt tgcacataat ccatatttgc
                                                                     - 1200
                                                                      1260
 tgttcaagtt aatctagaaa tttattcaat tctgtatgaa cacctggaag caaaatcata
 gtgcaaaaat acatttaagg tgtggtcaaa aataagtctt taattggtaa ataataagca
                                                                      1320
 ttaatttttt atagcctgta ttcacaattc tgcggtacct tattgtacct aagggattct
                                                                      1380
 aaaggtgttg tcactgtata aaacagaaag cactaggata caaatgaagc ttaattacta
                                                                      1440
                                                                      1500
 aaatgtaatt cttgacactc tttctataat tagcgttctt cacccccacc cccacccca
                                                                      1560
 cccccttat tttccttttg tctcctggtg attaggccaa agtctgggag taaggagagg
 attaggtact taggagcaaa gaaagaagta gcttggaact tttgagatga tccctaacat
                                                                      1620
 actgtactac ttgcttttac aatgtgttag cagaaaccag tgggttataa tgtagaatga
                                                                      1680
 tgtgctttct gcccaagtgg taattcatct tggtttgcta tgttaaaact gtaaatacaa
                                                                      1740
                                                                      1800
 1842
 aaaaaaaaa aaaaaaaaa aaaaaaaaaa aaaaaaanaa aa
<210> 91
<211> 1963
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (335)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1959)
<223> n equals a,t,g, or c
<400> 91
                                                                         60
 ggatcctcgc ggcggcggcg gtgcttacag cctgagaaga gcgtctcgcc cgggagcggc
                                                                        120
 ggcggccatc gagacccacc caaggcgcgt ccccctcggc ctcccagcgc tcccaagccg
                                                                        180
 cageggeege geceetteag ctageteget egetegetet getteeetge tgeeggetge
                                                                        240
 gcatggcktt ggcgttggcg gcgctggcgg cggtcgagcc gcctgcgcag ccggtaccag
                                                                        300
 cagttgcaga atgaagaaga gtctggagaa cctgaacagg ctgcaggtga tgctcctcca
 ccttacagca gcatttctgc agagagcgca gcatnatttt gactacaagg atgagtctgg
                                                                        360
 gtttccaaag cccccatctt acaatgtagc tacaacactg cccagttatg atgaagcgga
                                                                        420
```

gaggaccaag gctgaagcta ctatcccttt ggttcctggg agagatgagg attttgtggg

tcgggatgat	tttgatgatg	ctgaccagct	gaggatagga	aatgatggga	ttttcatgtt	540
aactttttc	atggcattcc	tctttaactg	gattgggttt	ttcctgtctt	tttgcctgac	60Ó
cacttcagct	gcaggaaggt	atggggccat	ttcaggattt	ggtctctctc	taattaaatg	660
gatcctgatt	gtcaggtttt	ccacctattt	ccctggatat	tttgatggtc	agtactggct	720
ctaataaata	ttccttgttt	taggetttet	cctgtttctc	agaggattta	tcaattatgc	780
aaaagttcgg	aagatgccag	aaactttctc	aaatctcccc	aggaccagag	ttctctttat	840
ttattaaaga	tgttttctgg	caaaggcctt	cctgcattta	tgaattctct	ctcaagaagc	900
aagagaacac	ctgcaggaag	tgaatcaaga	tgcagaacac	agaggaataa	tcacctgctt	960
taaaaaaata	aagtactgtt	gaaaagatca	tttctctcta	tttgttccta	ggtgtaaaat	1020
tttaatagtt	aatgcagaat	tctotaatca	ttgaatcatt	agtggttaat	gtttgaaaaa	1080
actettacaa	tcaagtctgt	gatgtattaa	taatgcctta	tatattgttt	gtagtcattt	1140
taagtagcat	gagccatgtc	cctgtagtcg	gtaggggca	gtcttgcttt	attcatcctc	1200
catctcaaaa	tgaacttgga	attaaatatt	gtaagatatg	tataatgctg	gccattttaa	1260
agggttttc	tcaaaagtta	aacttttgtt	atgactgtgt	ttttgcacat	aatccatatt	1320
tactattaa	gttaatctag	aaatttatto	aattctgtat	gaacacctgg	aagcaaaatc	1380
stantocaaa	aatacattta	addeddate	aaaaataaqt	ctttaattgg	taaataataa	1440
acaytycada	tttatagcct	gtattcacaa	trctgcggta	ccttattgta	cctaagggat	1500
tctasaggtg	ttgtcactgt	atasaacada	aagcactagg	atacaaatga	agcttaatta	1560
ctanaggtg	attcttgaca	ctctttctat	aattaccatt	cttcacccc	accccaccc	1620
ccaaaacyca	tattttcctt	ttatatata	atasttagac	caaagtctgg	gagtaaggag	1680
ccacccccc	acttaggagc	22222222	gtacttage	acttttgaga	tgatccctaa	1740
aggartaggr	tacttgcttt	tagaatgtat	tagcagaaac	cagtagatta	taatgtagaa	1800
tactetact	tctgcccaag	tactactes	tattaattta	ctatottass	actotaaata	1860
Egatgtgett	ttaataaata	tatattatat	acceptite	222222222	ааааааааа	1920
	aaaaaaaaa					1963
aaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa			
<210> 92 <211> 1487 <212> DNA <213> Homo s	sapiens					•
<220>						
<221> SITE						
<222> (1470))					
<223> n equa	als a,t,g, d	or c				
<220>						•
<221> SITE						
<222> (1487)						
<223> n equa	als a,t,g, o	or c				
400 00						
<400> 92				******	accaactaca	60 .
gcgaccgcgc	ccctttcagc	tagetegete	getegetetg	etteeetget	geeggeegeg	120
catggckwtg	gcgttggcgg	cgctggcggc	ggtcgagccg	geetgegeag	testestes	180
cagttgcaga	atgaagaaga	gtctggagaa	cctgaacagg	ctgcaggtga	cgcccccca	240
ccttacagca	gcatttctgc	agagagcgca	gtttccacc	tatttecttg	gatattttga	300
tggtcagtac	tggctctggt	gggtgttcct	tgttttaggc	tttctcctgt	ttettagagg	360
atttatcaat	tatgcaaaag	ttcggaagat	gccagaaact	ttctcaaatc	ccccaggac	
cagagttctc	tttatttatt	aaagatgttt	tctggcaaag	gccttcctgc	acctatgaat	420
tctctctcaa	gaagcaagag	aacacctgca	ggaagtgaat	caagatgcag	aacacagagg	480
aataatcacc	tgctttaaaa	aaataaagta	ctgttgaaaa	gatcatttct	ctctatttgt	540
tcctaggtgt	aaaattttaa	tagttaatgc	agaattctgt	aatcattgaa	ccactagtgg	600
ttaatgtttg	aaaaagctct	tgcaatcaag	tctgtgatgt	attaataatg	ccttatatat	660
tgtttgtagt	cattttaagt	agcatgagcc	atgtccctgt	agtcggtagg	gggcagtctt	720
gctttattca	tcctccatct	caaaatgaac	ttggaattaa	atattgtaag	atatgtataa	780
tgctggccat	tttaaagggg	ttttctcaaa	agttaaactt	ttgttatgac	cgcgcccccg	840
cocstastcc	atatttacta	ttcssattss	totagaaatt	tattcaattc	tgtatgaaca	900

cacataatcc atatttgctg ttcaagttaa tctagaaatt tattcaattc tgtatgaaca

PCT/US99/15849

```
cctggaagca aaatcatagt gcaaaaatac atttaaggtg tggtcaaaaa taagtcttta
                                                                        960
attggtaaat aataagcatt aattttttat agcctgtatt cacaattctg cggtacctta
                                                                       1020
ttgtacctaa gggattctaa aggtgttgtc actgtataaa acagaaagca ctaggataca
                                                                       1080
                                                                       1140
aatgaagett aattactaaa atgtaattet tgacactett tetataatta gegttettea
ccccaccc caccccacc cccttattt tccttttgtc tcctggtgat taggccaaag
                                                                       1200
                                                                       1260
 tctgggagta aggagaggat taggtactta ggagcaaaga aagaagtagc ttggaacttt
 tgagatgatc cctaacatac tgtactactt gcttttacaa tgtgttagca gaaaccagtg
                                                                       1320
 ggttataatg tagaatgatg tgctttctgc ccaagtggta attcatcttg gtttgctatg
                                                                       1380
 ttaaaactgt aaatacaaca gaacattaat aaatatctct tgtgtagcac ctttaaaaaa
                                                                       1440
 aaaaaaaaa aaaaaaaaa aaaaaaaaan cccggggggg ggccccn
                                                                       1487
<210> 93
<211> 1653
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (67)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (212)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1636)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1653)
<223> n equals a,t,g, or c
<400> 93
                                                                          60
 gegacegege cetteageta getegetege tegetetget tecetgetge eggetgegea
 tggcttnggc gttggcggcg ctggcggcgg ctcgagccgc ctgcgsagcc ggtaccagca
                                                                         120
                                                                         180
 gttgcagaat gaagaagagt ctggagaacc tgaacaggct gcaggtgatg ctcctccacc
 ttacagcagc atttctgcag agagcgcaca tnattttgac tacaaggatg agtctgggtt
                                                                         240
 tccaaagccc ccatcttaca atgtagctac aacactgccc agttatgatg aagcggagag
                                                                         300
 gaccaaggct gaagctacta tccctttggt tcctgggaga gatgaggatt ttgtgggtcg
                                                                         360
 ggatgatttt gatgatgctg accagctgag gataggaaat gatgggattt tcatgttaac
                                                                         420
                                                                         480
 ttttttcatg gcattcctct ttaactggat tgggtttttc ctgtcttttt gcctgaccac
                                                                         540
 ttcagctgca ggaaggtatg gggccatttc aggatttggt ctctctctaa ttaaatggat
                                                                         600
 cctgattgtc aggttttcca cctatttccc tgcatttatg aattctctct caagaagcaa
 gagaacacct gcaggaagtg aatcaagatg cagaacacag aggaataatc acctgcttta
                                                                         660
 aaaaaataaa gtactgttga aaagatcatt tctctctatt tgttcctagg tgtaaaattt
                                                                         720
 taatagttaa tgcagaattc tgtaatcatt gaatcattag tggttaatgt ttgaaaaagc
                                                                         780
 tcttgcaatc aagtctgtga tgtattaata atgccttata tattgtttgt agtcatttta
                                                                         840
 agtagcatga gccatgtccc tgtagtcggt agggggcagt cttgctttat tcatcctcca
                                                                         900
 tctcaaaatg aacttggaat taaatattgt aagatatgta taatgctggc cattttaaag
                                                                         960
 gggttttctc aaaagttaaa cttttgttat gactgtgttt ttgcacataa tccatatttg
                                                                        1020
 ctgttcaagt taatctagaa atttattcaa ttctgtatga acacctggaa gcaaaatcat
                                                                        1080
 agtgcaaaaa tacatttaag gtgtggtcaa aaataagtct ttaattggta aataataagc
                                                                        1140
 attaattttt tatagcctgt attcacaatt ctgcggtacc ttattgtacc taagggattc
                                                                        1200
```

PCT/US99/15849

1020

59

```
taaaggtgtt gtcactgtat aaaacagaaa gcactaggat acaaatgaag cttaattact
                                                                        1260
aaaatgtaat tottgacact otttotataa ttagogttot toaccoccac coccaccocc
                                                                        1320
accccctta ttttcctttt gtctcctggt gattaggcca aagtctggga gtaaggagag
                                                                        1380
gattaggtac ttaggagcaa agaaagaagt agcttggaac ttttgagatg atccctaaca
                                                                        1440
                                                                        1500
tactgtacta cttgctttta caatgtgtta gcagaaacca gtgggttata atgtagaatg
atgtgctttc tgcccaagtg gtaattcatc ttggtttgct atgttaaaac tgtaaataca
                                                                        1560
acagaacatt aataaatatc tcttgtgtag caccttttaw aaaaaaaaaa aaaaaaaaaa
                                                                        1620
                                                                        1653
 aaaaaaaaa aaaaancccg ggggggggcc ccn
<210> 94
<211> 1830
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (67)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (97)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (211)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1813)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1830)
<223> n equals a,t,g, or c
<400> 94
 gcgaccgcgc ccttcagcta gctcgctcgc tcgctctgct tccctgctgc cggctgcgca
                                                                          60
 tggcttnggc gttggcggcg ctggcggcgg tcgagcngcc tgcgsagccg gtaccagcag
                                                                         120
 ttgcagaatg aagaagagtc tggagaacct gaacaggctg caggtgatgc tcctccacct
                                                                         180
                                                                         240
 tacagcagca tttctgcaga gagcgcacat nattttgact acaaggatga gtctgggttt
 ccaaagcccc catcttacaa tgtagctaca acactgccca gttatgatga agcggagagg
                                                                         300
                                                                         360
 accaaggctg aagctactat ccctttggtt cctgggagag atgaggattt tgtgggtcgg
 gatgattttg atgatgctga ccagctgagg ataggaaatg atgggatttt catgttaact
                                                                         420
 tttttcatgg cattcctctt taactggatt gggtttttcc tgtctttttg cctgaccact
                                                                         480
 tcagctgcag gaaggtatgg ggccatttca ggatttggtc tctctctaat taaatggatc
                                                                         540
                                                                         600
 ctgattgtca ggttttccac ctatttccct ggatattttg atggtcagta ctggctctgg
 tgggtgttcc ttgttttagg ctttctcctg tttctcagag gatttatcaa ttatgcaaaa
                                                                         660
 gttcggaaga tgccagaaac tttctcaaat ctccccagga ccagagttct ctttatttat
                                                                         720
 taaagatgtt ttctggcaaa ggccttcctg catttatgaa ttctctctca agaagcaaga
                                                                         780
 gaacacctgc aggaagtgaa tcaagatgca gaacacagag gaataatcac ctgctttaaa
                                                                         840
 aaaataaagt actgttgaaa agatcatttc tctctatttg ttcctaggtg taaaatttta
                                                                         900
 atagttaatg cagaattctg taatcattga atcattagtg gttaatgttt gaaaaagctc
                                                                         960
```

ttgcaatcaa gtctgtgatg tattaataat gccttatata ttgtttgtag tcattttaag

PCT/US99/15849

			60			
			agageagtet	toctttattc	atcctccatc	1080
tagcatgagc	catgtccctg	cagicggtag	gaggeageee	atactaacca	ttttaaaggg	1140
tcaaaatgaa	cttggaatta	aacaccycaa	statatttt	gcacataatc	catatttgct	1200
gttttctcaa	aagttaaact	tttgttatga	ctgtgtttt	acctggaage	aaaatcatag	1260
gttcaagtta	atctagaaat	ctattcaatt	ctgtatgaat	actigguage	taataagcat	1320
tgcaaaaata	catttaaggt	gtggtcaaaa	acaagccccc	attetaceta	aggatteta	1380
taattttta	tagcctgtat	tcacaattct	geggtacett	accycaccca	taattactaa	1440
aaggtgttgt	cactgtataa	aacagaaagc	actaggatac	aaatgaaget	caaccaccaa	1500
aatgtaattc	ttgacactct	ttctataatt	agcgttcttc	accccaccc	ceacececae	1560
ccccttatt	ttccttttgt	ctcctggtga	ttaggccaaa	gtctgggagt	aaggagagga	1620
ttaggtactt	aggagcaaag	aaagaagtag	cttggaactt	ttgagatgat	ccctaacata	1680
ctgtactact	tgcttttaca	atgtgttagc	agaaaccagt	gggttataat	gtagaatgat	
atactttcta	cccaagtggt	aattcatctt	ggtttgctat	gttaaaactg	taaatacaac	1740
agaacattaa	taaatatctc	ttgtgtagca	ccttttaaaa	aaaaaaaaa	aaaaaaaaa	1800
	aancccgggg					1830
						•
<210> 95						
<211> 1134						
<212> DNA						
<213> Homo s	aniens					
(213) HOMO 2	aprono					
<400> 95						
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	gtcctcacac	aggraffcag	gaataccagg	atggcgtgcc	caagattcca	60
cccacctaca	ttacggtgga	aggedeedag	atgatgtcaa	gaatggcttc	tcatgggatc	120
acageetgta	ttcagctaaa	agacycagaa	angacetace	cagatactga	ttccttcaac	180
aaaattgtca	ttcagctaaa	gatgggggta	aagacccacc	ttatactagt	cagtggacat	240
actgtagcag	agatcactgg	gagcaaatat	ccagaacagg	acastagaa	ctttatatca	300
ctggacagct	gggatgttgg	gcagggtgcc	acggacgacg	geggeggage	tetacaacta	360
tgggaagcac	tctcacttat	taaagatctt	gggctgcgtc	caaayayyac	teagttacac	420
gtgctctgga	ctgcagaaga	acaaggtgga	gttggtgcct	ccagtatta	cttcttaccc	480
aaggtaaata	tttccaacta	cagtctggtg	atggagtctg	acgcaggaac	-ttotcaccc	540
actgggctgc	aattcactgg	cagtgaaaag	gccagggcat	catggaggag	gttatgaget	600
tgctgcagcc	cctcaatatc	actcaggtcc	tgagccatgg	agaagggaca	gacatcaact	660
tttggatcca	agctggagtg	cctggagcca	gtctacttga	tgacttatac	aagtatttt	720
tettecatea	ctcccacgga	gacaccatga	ctgtcatgga	tccaaagcag	atgaatgttg	
ctactactat	ttgggctgtt	gtttcttatg	ttgttgcaga	catggaagaa	atgetgeeta	780
ggtcctagaa	acagtaagaa	agaaacgttt	tcatgcttct	ggccaggaat	cctgggtctg	840
caactttgga	aaactcctct	tcacataaca	atttcatcca	attcatcttc	aaagcacaac	900
tctatttcat	actttctatt	attatctttc	ttgatacttt	ccaaattctc	tgcattctag	960
aaaaaggaat	cattctcccc	tccctcccac	cacatagaat	caacatatgg	tagggattac	1020
agtgggggca	tttctttata	tcacctctta	aaaacattgt	ttccacttta	aaagtaaaca	1080
cttaataaat	ttttggaaga	tctctgaaaa	aaaaaaaaa	aaaaaaaaa	aaaa	1134
						•
<210> 96						
<211> 1772						
<212> DNA						
<213> Homo	canione					
\213> Homo	Saprens					
<400> 96						
4002 30		~atcccade	caaatcccaa	acctatacct	gagcctgagc	60
Legaceacg	cyceegggag	gaccccage	cadaaactca	gaactataaa	accgctgggc	120
ctgagcctga	gccgagccgg	gageeggeeg	ttattaaast	taactcctta	ctcagectgt	180
ccccagcgat	ggcgaccccg	Lggggaggee		. daggeeeeeg	ctcagcctgt	240
cgtgcctggc	gctttccgtg	ctgctgctgg	cycactytta	gacyccycca	agaatttcga	300
ggatgtcaga	tgtaaatgta	tetgeetee	ctataaagaa	tataasaasa	catatttata	360
ataagaacat	atctcagaaa	gattgtgatt	gccttcatgt	g.ggagccc	atgcctgtgc	420
gggggcctga	tgtagaagca	tactgtctac	gctgtgaatg	caaatatgaa	gaaagaagct	480
ctgtcacaat	caaggttacc	attataattt	atctctccat	tttgggcctt	ctacttctgt	
acatggtata	tcttactctg	gttgagccca	tactgaagag	gegeetett	ggacatgcac	540

agttgataca	gagtgatgat	gatattgggg	atcaccagcc	ttttgcaaat	gcacacgatg	600
tgctagcccg	ctcccgcagt	cgagccaacg	tgctgaacaa	ggtagaatat	ggcacagcag	660
				tttgaccggc		720
				acaggcagac		780
				ttgatttcac		840
				tttttcttgt		900
atagagacat	ttttaaaagc	acacagetea	aagtcagcca	ataagtcttt	tcctatttgt	960
gacttttact	aataaaaata	aatctgcctg	taaaataaat	taaaaaatcc	tttacctgga	1020
				ttccaagata		1080
ttttattat	tattatttt	tatttattta	ttttaataaa	agaggggagg	gatgcctggg	1140
				acttttgtaa		1200
				cagcctcatc		1260
				atctgctgtg		1320
ataataaaca	gatetaga	tacctaataa	cttttcacaa	aaagcagatt	ttcttcatqt	1380
actgtaatgt	ctgatggaat	gcatcctaga	acaaactggc	catttgctag	tttactctaa	1440
acceptgacge	tagtetteet	atatataata	ttactcatct	tctagtacct	ttaaggacaa	1500
atcctaacca	cttqqacact	tacaataaaa	aaattttatt	ttaaacccaa	acctccctaa	1560
attestasta	tatacacatt	totcaccatt	tccaatcata	gtgagaggca	actatttaaa	1620
ctccaatata	tacacatate	aactagggct	agaattataa	gtgcctcttc	tgaaaggtct	1680
ascettatt	grataactor	cttttttt	tectetttaa	aatgtaacaa	taaaaataat	1740
				aa ag aaaaaa		1772
·	ccaaaaaaaa	aaaaaaaaa	~~			

<210> 97 <211> 2381 <212> DNA

<213> Homo sapiens

<400> 97

60 ccacgcgtcc cgcaaggcca gttctagtgt agagagaaaa aggagccggc agcggctctt acgcgtcccg gggctgcgcg ccactctctc ggccggtaac gcggtgcttt gcggctgtcg 120 tcaagcgcgg cgttgggccg gcgggcgggg gctgaggggc tgccatggcg gcggcgggcc 180 ggctcccgag ctcctgggcc ctcttctcgc cgctcctcgc agggcttgca ctactgggag 240 tcgggccggt cccagcgcgg gcgctgcaca acgtcacggc cgagctcttt ggggccgagg 300 360 cctggggcac ccttgcggct ttcggggacc tcaactccga caagcagacg gatctcttcg 420 tgctgcggga aagaaatgac ttaatcgtct ttttggcaga ccagaatgca ccctatttta 480 aacccaaagt aaaggtatct ttcaagaatc acagtgcatt gataacaagt gtagtccctg 540 gggattatga tggagattct caaatggatg tccttctgac atatcttccc aaaaattatg 600 ccaagagtga attaggagct gttatcttct ggggacaaaa tcaaacatta gatcctaaca atatgaccat actcaatagg acttttcaag atgagccact aattatggat ttcaatggtg 660 atctaattcc tgatattttt ggtatcacaa atgaatccaa ccagccacag atactattag 720 780 gagggaattt atcatggcat ccagcattga ccactacaag taaaatgcga attccacatt 840 ctcatgcatt tattgatctg actgaagatt ttacagcaga tttattcctg acgacattga atgccaccac tagtaccttc cagtttgaaa tatgggaaaa tttggatgga aacttctctg 900 960 tcagtactat attggaaaaa cctcaaaata tgatggtggt tggacagtca gcatttgcag 1020 actttgatgg agatggacac atggatcatt tactgccagg ctgtgaagat aaaaattgcc aaaagagtac catctactta gtgagatctg ggatgaagca gtgggttcca gtcctacaag 1080 atttcagcaa taagggcaca ctctggggct ttgtgccatt tgtggatgaa cagcaaccaa 1140 1200 ctgaaatacc aattccaatt accettcata ttggagacta caatatggat ggctatccag 1260 acgctctggt catactaaag aacacatctg gaagcaacca gcaggccttt ttactggaga 1320 acgtcccttg taataatgca agctgtgaag aggcgcgtcg aatgtttaaa gtctactggg 1380 agctgacaga cctaaatcaa attaaggatg ccatggttgc caccttcttt gacatttacg aagatggaat cttggacatt gtagtgctaa gtaaaggata tacaaagaat gattttgcca 1440 1500 ttcatacact aaaaaataac tttgaagcag atgcttattt tgttaaagtt attgttctta gtggtctgtg ttctaatgac tgtcctcgta gataacaccc tttggagtga atcaacctgg 1560 accttatate atgtatacaa etgtagatge aaatgggtat etgaaaaatg gateagetgg 1620 ccaactcagc caatccgcac atttagctct ccaactacca tacaacgtgc ttggtttagg 1680 tcggagcgca aattttcttg accatctcta cgttggtatt ccccgtccat ctggagaaaa 1740

1920 1955

atctatacga	aaacaagagt	ggactgcaat	cattccaaat	tcccagctaa	ttgtcattcc	1800
ataccctcac	aatotogogo	gaagttggag	tgccaaactg	tatcttacac	caagtaatat	1860
tgttctgctt	actoctatao	ctctcatcgg	tatctatatt	ttcatcttgg	caataattgg	1920
cattttacat	taacaaaaa	agaaagcaga	tgatagagaa	aaacgacaag	aagcccaccg	1980
gtttcatttt	astactatat	gacttgcctt	taatattaca	taatggaatg	gctgttcact	2040
tgattagttg	gatgetatge	tetagettaa	aaaaataggg	gagattaaat	attatttata	2100
aatgatgtat	adacacaaac	ttattagaaa	gtattcaaat	aaatatggtt	tgaatatgtc	2160
acaaggtctt	cccatygtaa	ggactttgta	tataaaaatt	tagattetet	attctgtagt	2220
gctgtacatt	tttattatt	tatageatat	attacatata	ctccagtgtt	tgtgtattta	2280
taatcttatt	tragtatat	tastaassa	agttgtgtaa	ataaaaataa	ttaaatgagc	2340
aggaaaaaaa	cgcatcatga	22222222	2222222222	a	-	2381
aggaaaaaaa	aaaaaaaaa	aaaaaaaaa	444444	_		
-010- 00						
<210> 98						
<211> 1955						
<212> DNA	aniona					
<213> Homo s	aptens					
<400> 98					•	
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	ccatccctat	atttactacc	atoctettee	ttttctccat	ggctacactg	60
ggcacgagcg	coattooctge	ccctagaga	attecteggg	cgctaccaga	tgaagcagct	120
ttgaggacca	taasastaas	ecctagages	agtacagtac	cccagggcca	gcgaccaccg	180
ttcatagaaa	rgagataga	getaecaac	cagattgtga	aactgaaata	ctgttacaca	240
ccccgcacca	tagaatttt	cogggette	cattgcagca	tetataacaa	ctgtgtggag	300
tgcaagatct	teeggeetee	ataggtaga	aattatata	gaaagaggaa	ctaccgctac	360
egettegace	accaetyeee	tetetecete	ctcacaatct	atotcttcoc	cttcaacatc	420
ttetacetet	catecute	tttassastt	agettettag	agacattgaa	aggaaactcc	480
gtctatgtgg	ccccaaacc	tantttactt	ctttacactc	taatccatca	tgggactgac	540
tggaactgtt	ctagaagtee	teactigett	CCECACACC	aatgaaagac	atcaaaggat	600
tggatttcat	actitecteg	changesta	ccagacaace	toocaatatt	gtgaagaact	660
catggacagg	gaagaatege	gcccayaacc	ccacageca	ggatcgaagg	ggtattttgc	720
gctgtgaagt	getgtgtge		ctagtgtgtt	gaccagtage	agcctcttgc	780
cactggagga	aagtggaagt	cgacetecca	gcacccaaga	gatacagaa	agcctcttgc gacagcagca	840
cacagagccc	agececcaca	gaacacctga	acceaacga	acaddaggca	gacagcagca gctgaagctg	900
ctcccgaaga	gatgccacct	ccagageeee	cagagecace	actggaggot	gctgaagctg atgagagatt	960
agaagtagcc	tatctatgga	agagacttt	gcccgcgccc	tatacattt	atgagagatt aactgtttt	1020
tcaggtgaga	agttaaacct	gagacagaga	gcaagtaagt	tactacaaac	aactgttttt tttttaaat	1080
ctttggtctt	tagtcaccca	gttgcacact	ggcaccccc	tastasctaa	ttttttaaat	1140
ttctgaactc	aaggcagtgg	cagaagatgt	cagucaccic	egacaacces	aaaaatgggt tggacccct	1200
ctcttgggcc	ctggcactgg	tteteeatgg	tagatasat	agggttett	tggacccct tggggctaaa	1260
ctcttccctc	cagatcccag	ccctcctgct	cggggtcact	ggcccacc	tggggctaaa	1320
agttttcgag	actggctcaa	atcctcccaa	getgetgeac	graceter	cagaggcagt	1380
cacagagacc	tctggccagg	ggatcctaac	tgggttellg	gggccccag	gactgaagag	1440
gagggagagt	ggggtcagaa	gattctcctg	gccaccaagt	, godaycateg	cccacaaatc	1500
cttttaggaa	tgggacaggt	accttccact	agttgtattt	. accaytytay	cttctccttt	1560
gtctcccatc	cactctgaca	ccttaagccc	dactettte	: ccactagata	tatgtaagta	1620
gttgtagtag	agataataat	tgacatttct	cgtagactac	ccagaaactt	ttttaatacc	1680
tgtgccattc	tcaataagaa	tttatgagat	. gccagcggca	Laycoccida	cactctctgt	1740
ctcatctctc	ctcctttctc	attagcccct	. tttaatttgt		gactcctgct	1800
cccattagga	gcaggaatgg	cagtaataaa	agtctgcact	. ctggtcattt	cttttcctca	1860
gaggaagcct	gagtgctcac	ttaaacacta	tcccctcaga	ctccctgtgt	gaggcctgca tcctctc	1920
		~~~===	. aacacaaaa	i aacicteett		1700

tccccgatg taccctcaaa aaaaaaaaa aaaaa

<210> 99

<211> 1958

<212> DNA

<213> Homo sapiens

<400> 99					aaaacaacat	60
ccacgcgtcc	ggggcgttçc	tggtcgtgag	aggggagccc	caggggaget	ggggcagcac	120
gactggggtg	ataaatggcc	ggaaatttgg	cgtggccaca	cccaacacca	gcgcgacgca	180
ggaggcacac	tccggggtca	gcagcatcca	cagcagcatc	egecacyccc	cagcaaacgc	240
ggggcctctg	atgcgggtgc	tcgtggtcac	categeeeee	acctactggg	aggagagtea	300
agagagtggg	gaagccctga	atggccactc	tctgactggg	ggcaagttcc	ggcagagtca.	360
cacgtggagt	ttgctacagg	gagetgetea	cgatgaccca	gtggcccggg	gcccggaccc	420
cgatggcctc	ctgctcctcg	acgtggtggt	caatggcgtt	gtccccggac	gageetgget	480
gacgcagatc	ttcaagtgca	ggactttgaa	gaagcactac	gtgcaaacaa	gggeetggee	540
agctgttcgt	gggctccaca	cagcgcttct	tccagggcgg	ceteeeeteg	atacaacac	600
gcaaccacag	catccagtac	aacgcggccc	ggggcccca	geeecagetg	graceacc	660
tgcgggcctc	agctatcagc	teggeetttg	acceagagge	egaggeeerg	tttgaggtgg	720
tcgctacagc	cctgcaggcg	gaggagaacg	aggreggerg	atagastast	cacctetece	780
actcccaggg	agcgttttgt	gtggatgtgg	acgagtgtgc	gtgggatget	accetteget	840
gagagggaca	gcgctgtgtg	aacctgctcg	ggtectaccg	etgeeteeee	gactgtgggt	900
ctggcttccg	ggtggctgat	ggggccggct	grgaaaargr	ggacgaatge	ccggaagggg	960
ttggacgact	gtcactacaa	ccagctctgc	gagaacaccc	caggeggtea	cogcogcage	1020
tgccccaggg	gttaccggat	gcagggcccc	agcetgeect	geetagatgt	caatgagtgc	1080
ctgcagctgc	ccaaggcctg	cgcctaccag	tgccacaacc	cccagggcag	ctaccyctyc	1140
ctgtgccccc	caggccagac	cctccttcgc	gacggcaagg	cetgeacete	actggagcgg	1200
aatggacaaa	atgtgaccac	cgtcagccac	cgaggccctc	cattgeectg	getgeggete	1260
tgggcctcga	tccccggtac	ctcctaccac	gcctgggtct	eteteegtee	gggtcccatg	1320
gccctgagca	gtgtgggccg	ggcctggtgc	ceteetggtt	caccaggea	gaacggagcc	1380
tgcacagacc	ttgacgagtg	ccgcgtgagg	aacctgtgtc	ageaegeeeg	ccgcaacacc	1440
gagggcagct	accagtgcct	gtgccccgcc	ggctaccgtc	tgetececag	cgggaagaac	1500
tgccaggaca	tcaacgagtg	cgaggaggag	agcatcgagt	gtggacccgg	ctagatgtgt	1560
ttcaacaccc	gtggcagcta	ccagtgtgtg	gacacaccct	geeetgeeac	acacacacac	1620
ggccccagcc	ctgggacgtg	cttccggcgc	tgctcgcagg	actgcggcac	gggcggcccc	1680
tctacgctgc	agtaccggct	gctgccgctg	cccctgggcg	tgegegeeea	ccacgacgtg	1740
gcccgcctca	ccgccttctc	cgaggtcggc	gtccccgcca	accgcaccga	gctcagcatg	1800
ctggagcccg	acccccgcag	ccccttcgcg	ctgcgtccgc	tgcgcgcggg	ccttggcgcg	1860
gtctacaccc	gtcgcgcgct	cacccgcgcc	ggcctctacc	ggctcaccgt	gcgtgctgcg	1920
gcaccgcgcc	accaaagcgt	cttcgtcttg	ctcatcgccg	tgtececcia	cccctactaa	1958
acgggagagg	gcattggcgg	ccgctctaga	ggatccct			1936
<210> 100						•
<211> 2444						
<212> DNA	•					
<213> Homo	sapiens					
-400- 100						
<400> 100	ctggcacgag	022502220	gttaatctct	ttaactaaac	ctacagatga	60
ccaegecaag	cuggeacyag	aggttcatgo	tttaaatato	cttagagcat	tgttcagaga	120
tagagaga	acagtecece	ttattcctta	tattactast	ggagctaagg	ctgcaattct	180
cacgegeetg	ggagaaaaca	caccecca	apattcatcc	acacttctct	ttagtgcctt	240
gggtttaca	etateggtet	ttaaaaaaaa	aaaccatcaa	cattccaaaa	caaatagaat	300
gattacaaya	acciciggag	ctadaagggc	agaactctat	cettttette	tcaaacagtt	360
gacagggaga	gagtttttt	tagagagag	tataaaaaa	ccaaatcgtc	atccaagcat	420
ggaaactgta	gecaatacag	tagacagega	ctacgggagaa	ccaaraata	gtacttcttc	480
gittetetta	atacasast	ttattaatt	cattatrarr	tataatcact	cacctgtcta	540
egeceteage	acgggacccc	stostasstt	aatcccattt	gttatgatag	atcacattcc	600
ccactcccgt	gaaatggcag	tatacacac	coccarcter	actoaccant	gtttccgggc	660
caacaccatt	cgaactctgt	nattataaa	attttccat	ttattacasa	cctactcaga	720
aaaaccacat	ccatggggac	taagagetaa	acceccac	gactgacate	actotttota	780
ctccaaaaca	cggaacgaat	ceagactice	agcacyayct	tttaataaco	actgtttgta agagctgtat	840
ccaaagccaa	actctggctg	gccaagaggc	acctcaccac	atctgcaaag	agagctgtat	900
acactgatat	cccctccca	Ligacitget	gccccaacag	accegeadag	gacaaccagc	3.00

64

cagttctgga	gagtcttggc	ttctgggaag	aaattcaaag	ggaattatct	caggatcaga	960
agctgataac	gggattccct	tgggccttca	aggtgccagg	cctgccccag	tacctccaga	1020
gcctcaccag	actagccatt	gctgcagtgt	gggccgcggc	agccaagagt	ggagagcggg	1080
agacgaatgt	ccccatctct	ttctctcagc	tgttagaatc	tgccttccct	gaagtgcgct	1140
cactaacact	ggaagccctc	ttggaaaagt	tcttagcagc	agactctgga	cttggagaga	1200
agggcgtgcc	acccttgctg	tgcaacatgg	gagagaagtt	cttattgttg	gccatgaagg	1260
aaaatcaccc	agaatgcttc	tgcaagatac	tgaaaattct	acactgcatg	gaccctggtg	1320
agtggcttcc	ccagacggag	cactgtgtcc	atctgacccc	aaaggagttc	ttgatctgga	1380
cgatggatat	tgcttccaat	gaaagatctg	aaattcagag	tgtagctctg	agacttgctt	1440
ccaaagtcat	ttcccaccac	atgcagacat	gtgtggagaa	cagggaattg	atagctgctg	1500
agctgaagca	gtgggttcag	ctggtcatct	tgtcatgtga	agaccatctt	cctacagagt	1560
ctaggctggc	cgtcgttgaa	gtcctcacca	gtactacacc	acttttcctc	accaaccccc	1620
atcctattct	tgagttgcag	gatacacttg	ctctctggaa	gtgtgtcctt	acccttctgc	1680
agagtgagga	gcaagctgtt	agagatgcag	ccacggaaac	cgtgacaact	gccatgtcac	1740
aagaaaatac	ctgccagtca	acagagtttg	ccttctgcca	ggtggatgcc	tccatcgctc	1800
tggccctggc	cctggccgtc	ctgtgtgatc	tgctccagca	gtgggaccag	ttggcccctg	1860
gactgcccat	cctgctggga	tggctgttgg	gagagagtga	tgacctcgtg	gcctgtgtgg	1920
agagcatgca	tcaggtggaa	gaagactacc	tgtttgaaaa	agcagaagtc	aacttttggg	1980
ccgagaccct	gatctttgtg	aaatacctct	gcaagcacct	cttctgtctc	ctctcaaagt	2040
ccggctggcg	tcccccaagc	cctgagatgc	tctgtcacct	tcaaaggatg	gtgtcagagc	2100
agtgccacct	cctgtctcag	ttcttcagag	agcttccacc	agctgctgag	tttgtgaaga	2160
cagtggagtt	cacaagacta	cgcattcaag	aggaaaggac	tttggcttgc	ttgaggctgc	2220
tggccttttt	ggaaggaaag	gaaggggaag	acaccctagt	tctcagtgtt	tgggactctt	2280
atgcagaatc	gaggcagtta	actcttccaa	gaacagaagc	ggcatgttga	agaaaatctg	2340
ggggattggg	atgggggtat	gtgtggattt	ttcctccact	aaatctgcag	gaaacatgtt	2400
gaacataaat	tcaaaaattt	tatcccaaaa	aaaaaaaaa	aaaa		2444

<210> 101 <211> 2709

<212> DNA

<213> Homo sapiens

<400> 101

4002 101						
ggcacgagat	ttcctacagg	tgaaacgcca	tcattaggat	tcactgtaac	gttagtgcta	60
ttaaactcac	tagcattttt	attaatggcc	gttatctaca	ctaagctata	ctgcaacttg	120
gaaaaagagg	acctctcaga	aaactcacaa	tctagcatga	ttaagcatgt	cgcttggcta	180
atcttcacca	attgcatctt	tttctgccct	gtggcgtttt	tttcatttgc	accattgatc	240
actgcaatct	ctatcagccc	cgaaataatg	aagtctgtta	ctctgatatt	ttttccatgc	300
ctgcttgcct	gaatccagtc	ctgtatgttt	tcttcaaccc	aaagtttaaa	gaagactgga	360
agttactgaa	gcgacgtgtt	accaagaaaa	gtggatcagt	ttcagtttcc	atcagtagcc	420
aaggtggttg	tctggaacag	gatttctact	acgactgtgg	catgtactca	catttgcagg	480
gcaacctgac	tgtttgcgac	tgctgcgaat	cgtttcttt	aacaaagcca	gtatcatgca	540
aacacttgat	aaaatcacac	agctgtcctg	cattggcagt	ggcttcttgc	caaagacctg	600
agggctactg	gtccgactgt	ggcacacatt	cggcccactc	tgattatgca	gatgaagaag	660
attcctttgt	ctcagacagt	tctgaccagg	tgcaggcctg	tggacgagcc	tgcttctacc	720
agagtagagg	attccctttg	gtgcgctatg	cttacaatct	accaagagtt	aaagactgaa	780
ctactgtgtg	tgtaaccgtt	tcccccgtca	accaaaatca	gtgtttatag	agtgaaccct	840
attctcatct	ttcatctggg	aagcacttct	gtaatcactg	cctggtgtca	cttagaagaa	900
ggagaggtgg	cagtttattt	ctcaaaccag	tcattttcaa	agaacaggtg	cctaaattat	960
aaattggtga	aaaatgcaat	gtccaagcaa	tgtatgatct	gtttgaaaca	aatatatgac	1020
ttgaaaagga	tcttaggtgt	agtagagcaa	tataatgtta	gttttttctg	atccataaga	1080
agcaaattta	tacctatttg	tgtattaagc	acaagataaa	gaacagctgt	taatatttt	1140
taaaaattct	attttaaaa	tgtgattttc	tataactgaa	gaaaaatatc	ttgctaattt	1200
tacctaatgt	ttcatccttt	aatctcagga	caacttactg	cagggccaaa	aaagggactg	1260
tcccagctag	acctgtgaga	gtatacatag	gcattacttt	attatgtttt	cacttgccat	1320
ccttgacata	agagaactat	aaattttgtt	taagcaattt	ataaatctaa	aacctgaaga	1380
tgtttttaaa	acaatattaa	cagctgttag	gttaaaaaaa	tagctggaca	tttgttttca	1440

```
gtcattatac attgctttgg tccaatcagt aattttttct taagtgtttt gtgattacac
                                                                      1500
tactagaaaa aaagtaaaag gctaattgct gtgtgggttt agtcgatttg gctaaactac
                                                                      1560
taactaatgt gggggtttaa tagtatctga gggatttggt ggcttcatgt aatgttctca
                                                                      1620
ttaatgaata cttcctaata tcgttggctc tactaatatt ttccaatttg ctgggatgtc
                                                                      1680
                                                                      1740
acctagcaat agcttggatt atatagaaag taaactgtgg tcaatacttg catttaatta
gacgaaacgg ggagtaatta tgacacgaag tacttaatgt ttatttctta gtgagctgga
                                                                      1800
                                                                      1860
ttatcttgaa cctgtgctat taaatggaaa tttccataca tcttccccat actattttt
                                                                      1920
ataaaagagc ctattcaata gctcagaggt tgaactctgg ttaaacaaga taatatgtta
                                                                      1980
ttaataaaaa tagaagaaga aagaataaag cttagtcctg tgtcttttaa aaattaaaaa
ttttacttga ttccccatct atgggcttta gacctattac tgggtggagt cttaaagtta
                                                                      2040
                                                                      2100
taattgttca atatgttttt tgaacagtgt gctaaatcaa tagcaaaccc actgccatat
tagttattct gaatatacta aaaaaatcca gctagattgc agtttaataa ttaaactgta
                                                                      2160
catactgtgc atataatgaa tttttatctt atgtaaatta tttttagaac acaagttggg
                                                                      2220
aaatgtggct tctgttcatt tcgtttaatt aaagctacct cctaaactat agtggctgcc
                                                                      2280
                                                                      2340
agtagcagac tgttaaattg tggtttatat actttttgca ttgtaaatag tctttgttgt
                                                                      2400
acattgtcag tgtaataaaa acagaatctt tgtatatcaa aatcatgtag tttgtataaa
atgtgggaag gatttattta cagtgtgttg taattttgta aggccaacta tttacaagtt
                                                                      2460
ttaaaaattg ctatcatgta tatttacaca tctgataaat attaaatcat aacttggtaa
                                                                      2520
gaaactccta attaaaaggt tttttccaaa attcaggtta ttgaaaactt ttcattttat
                                                                      2580
tcatttaaaa actagaataa cagatatata aaagtgttaa tctttgtgct atatggtatg
                                                                      2640
2700
                                                                      2709
aaaaaaaa
<210> 102
<211> 1722
<212> DNA
<213> Hómo sapiens
<220>
<221> SITE
<222> (401)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (695)
<223> n equals a,t,g, or c
gggaccgcgc tgtcctgctg tcaccaagag ctggagacac catctcccac cgagagtcat
                                                                        60
                                                                       120
ggcccattg gccctgcacc tcctcgtcct cgtccccatc ctcctcagcc tggtggcctc
ccaggactgg aaggctgaac gcagccaaga ccccttcgag aaatgcatgc aggatcctga
                                                                       180
                                                                       240
ctatgagcag ctgctcaagg tcaccatcct ggaggcagat aacaggatcg ggggccgcat
                                                                       300
cttcacctac cgggaccaga wyacgggctg gattggggag ctgggagcca tgcgcatgcc
cagctctcac aggatcctcc acaagctctg ccagggcctg gggctcaacc tgaccaagtt
                                                                       360
cacccagtac gacaagaaca cgtggacgga ggtgcacgaa ntgaagctgc gcaactatgt
                                                                       420
                                                                       480
ggtggagaag gtgcccgaga agctgggcta cgccttgcgt ccccaggaaa agggccactc
gcccgaagac atctaccaga tggctctcaa ccaggccctc aaagacctca aggcactggg
                                                                       540
ctgcagaaag gcgatgaaga agtttgaaag gcacacgctc ttggaatatc ttctcgggga
                                                                       600
ggggaacctg agccggccgg ccgtgcagct tctgggagac gtgatgtccg aggatggctt
                                                                       660
                                                                       720
 cttctatctc agcttcgccg aggccctccg ggccnacagc tgcctcagcg acagactcca
gtacagccgc atcgtgggtg gctgggacct gctgccgcgc gcgctgctga gctcgctgtc
                                                                       780
                                                                       840
 cgggcttgtg ctgttgaacg cgcccgtggt ggcgatgacc cagggaccgc acgatgtgca
cgtgcagatc gagacctctc ccccggcgcg gaatctgaag gtgctgaagg ccgacgtggt
                                                                       900
                                                                       960
gctgctgacg gcgagcggac cggcggtgaa gcgcatcacc ttctcgccgc gctgccccgc
cacatgcagg aggcgctgcg gaggctgcac tacgtgccgg ccaccaaggt gttcctaagc
                                                                      1020
 ttccgcaggc ccttctggcg cgaggagcac attgaaggcg gccactcaaa caccgatcgc
                                                                      1080
```

WO 00/04140

PCT/US99/15849

66												
ccgtcgcgca tgattttcta cccgccgcg cgcgagggcg acgtggtcgg acgcggcggc agcgttcgcc ggcttgagcc gcgctcgacg acgtggcgc attgcacggg cctgtcgtgc ggcgtcgtca agcgttggc ggaggaccag cacagccagg ccggcgctct ggcaaaccga aaaggatgac tggacggtcc gccggcgagc acaccgccta cccgcacggc tgggtggaga cgcgccgca tcaagatcaa cagccggaag gggcctgcat gggcacgcat ctgacatgga gggcagggg catgtgcatg catgacctgg caaaggaaga aggcagccac cctccagtcc aacacgaccc acacgaggac ctcgcattaa agtattttcg aaaaaaaaaa aaaaaaaaaa aaaaaagggcgg	gggaagaggc gttgcgcttg gccagctctg ggacggcacc gtggctttgt ggtacagmcg cttatggccg catctacttt cggcggtcaa gtcggcgctg cggacacggc cagccccgag gggtggccag cagcccctcg aaggccagtt atctctcaa gaaaaaaaaa aaaaaaaaa											
<210> 103												
<211> 106												
<212> PRT												
<213> Homo sapiens												
<220>												
<221> SITE												
<pre><222> (14) <223> Xaa equals any of the naturally occurr</pre>	ing L-amino acids											
very new equals any of the medically occur												
<220>												
<221> SITE												
<222> (29)												
<223> Xaa equals any of the naturally occurr	ing L-amino acids											
<400> 103												
Met Gly Ser Leu Ser Gly Cys Ala Leu Pro Phe	Cys Leu Xaa Val Phe											
1 5 10	15											
Phe Leu Thr Val Ser Pro Ser Ala Val Gly Leu												
20 25	30											
Gly Pro Leu Gln Thr Leu Phe Ala Trp Val Ser	Pro Val Glu Ala Ala											
35 40	45											
Glu Gln Gln Arg Leu Leu Pro Val Leu Ser Ser												
50 55	60											
Clar Clar Man Care Clar Man Dune Ale Aver Ale Lou	rou mur Clu Val Sor											
Glu Gly Thr Cys Gln Met Pro Ala Arg Ala Leu 65 70 75	Bed Tyl Gld val Sel											
70 73	00											
Val Gly Pro Tyr Trp Glu Ile Pro Pro Ser Gln	Asp Thr Arg Arg Ser											
85 90	95											
Gly Thr Tyr Leu Arg Arg Gln Ser Asp Pro												
100 105												
<210> 104												
<211> 86												
<212> PRT												
<213> Homo sapiens												

<400> 104

Met Thr Leu Pro Ser Arg Ala Leu Ala Ser Leu Gly Val Gly Val Trp

67 15 10 Gly Met Leu Arg Leu Asn Gln Val Thr Val Ser Cys Gly Gly Ser Arg 20 Trp Ser Ser Arg Val Ala Leu Gly Ala Phe Ser Trp Val Cys Gly Val Ala Leu Val Leu Gln Pro Ser Gly Gly Gly Leu Gly Leu Thr Ser Pro Ser Glu Gly Cys Trp Glu Gly Glu Leu Ala Leu Ala Val Leu Arg Ala Pro Gly Gly Ser Pro Ser <210> 105 <211> 302 <212> PRT <213> Homo sapiens <400> 105 Met Ala Arg Ala Arg Gly Ser Pro Cys Pro Pro Leu Pro Pro Gly Arg Met Ser Trp Pro His Gly Ala Leu Leu Phe Leu Trp Leu Phe Ser Pro Pro Leu Gly Ala Gly Gly Gly Val Ala Val Thr Ser Ala Ala Gly 40 Gly Gly Ser Pro Pro Ala Thr Ser Cys Pro Val Ala Cys Ser Cys Ser 50 Asn Gln Ala Ser Arg Val Ile Cys Thr Arg Arg Asp Leu Ala Glu Val Pro Ala Ser Ile Pro Val Asn Thr Arg Tyr Leu Asn Leu Gln Glu Asn 90 Gly Ile Gln Val Ile Arg Thr Asp Thr Phe Lys His Leu Arg His Leu 100 Glu Ile Leu Gln Leu Ser Lys Asn Leu Val Arg Lys Ile Glu Val Gly 120 Ala Phe Asn Gly Leu Pro Ser Leu Asn Thr Leu Glu Leu Phe Asp Asn 130 135 Arg Leu Thr Thr Val Pro Thr Gln Ala Phe Glu Tyr Leu Ser Lys Leu 150 155 Arg Glu Leu Trp Leu Arg Asn Asn Pro Ile Glu Ser Ile Pro Ser Tyr 165

Ala Phe Asn Arg Val Pro Ser Leu Arg Arg Leu Asp Leu Gly Glu Leu

68 185 190 180 Lys Arg Leu Glu Tyr Ile Ser Glu Ala Ala Phe Glu Gly Leu Val Asn 200 195 Leu Arg Tyr Leu Asn Leu Gly Met Cys Asn Leu Lys Asp Ile Pro Asn Leu Thr Ala Leu Val Arg Leu Glu Glu Leu Glu Leu Ser Gly Asn Arg 225 230 Leu Asp Leu Ile Arg Pro Gly Ser Phe Gln Gly Leu Thr Ser Leu Arg Lys Leu Trp Leu Met His Ala Gln Val Ala Thr Ile Glu Arg Asn Ala Phe Asp Asp Leu Lys Ser Leu Glu Glu Leu Asn Leu Ser His Asn Asn 280 275 Leu Met Ser Leu Pro His Asp Leu Phe Thr Pro Leu His Arg 295 300 <210> 106 <211> 56 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (56) <223> Xaa equals stop translation <400> 106 Met Pro Ser Ser Trp Leu Pro Gly Cys Phe Val Leu Leu Cys Leu Val 5 Ala Val Gly Cys Gln Leu Arg Glu Trp Gly Val Gly Val Ser Ala Val Gly Leu Leu Ala Leu Pro His Leu Gln Val Leu Gly Met Arg Gly 40 Arg Gly Leu Ile Ser Gly Gly Xaa <210> 107 <211> 189 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (94) <223> Xaa equals any of the naturally occurring L-amino acids

WO 00/04140

69

<400> 107

Met Ala Leu Leu Ser Arg Pro Ala Leu Thr Leu Leu Leu Leu Met 10

Ala Ala Val Val Arg Cys Gln Glu Gln Ala Gln Thr Thr Asp Trp Arg

Ala Thr Leu Lys Thr Ile Arg Asn Gly Val His Lys Ile Asp Thr Tyr

Leu Asn Ala Ala Leu Asp Leu Leu Gly Gly Glu Asp Gly Leu Cys Gln 50

Tyr Lys Cys Ser Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys

Pro Ser Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Xaa His Leu

Asn Ile Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg

Cys Tyr Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe 120

Gln Tyr Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly 135

Leu Thr Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe 155

Asp Ser Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg

Ala Ala Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu 180 185

<210> 108

<211> 61

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (61)

<223> Xaa equals stop translation

<400> 108

Met Gly Asn Cys Gln Ala Gly His Asn Leu His Leu Cys Leu Ala His 10

His Pro Pro Leu Val Cys Ala Thr Leu Ile Leu Leu Leu Gly Leu 20

Ser Gly Leu Gly Leu Gly Ser Phe Leu Leu Thr His Arg Thr Gly Leu 40

70

Arg Thr Leu Thr Ser Pro Arg Thr Gly Ser Leu Phe Xaa 50 55 60

<210> 109

<211> 128

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (90)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 109

Met Arg Leu Glu Ser Leu Cys His Leu Cys Leu Ala Cys Leu Phe Phe 1 5 10 15

Arg Leu Pro Ala Thr Arg Thr Val Tyr Cys Met Asn Glu Ala Glu Ile 20 25 30

Val Asp Val Ala Leu Gly Ile Leu Ile Glu Ser Arg Lys Gln Xaa Lys
35 40 45

Ala Cys Glu Gln Pro Ala Leu Ala Gly Ala Asp Asn Pro Glu His Ser

Pro Pro Cys Ser Val Ser Pro His Thr Ser Ser Gly Ser Ser Ser Glu 65 70 75 80

Glu Glu Asp Ser Gly Lys Gln Ala Leu Xaa Pro Gly Leu Ser Pro Ser 85 90 95

Gln Arg Pro Gly Gly Ser Ser Ser Ala Cys Ser Arg Ser Pro Glu Glu 100 105 110

Glu Glu Glu Asp Val Leu Lys Tyr Val Arg Glu Ile Phe Phe Ser 115 120 125

<210> 110

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (50)

<223> Xaa equals any of the naturally occurring L-amino acids

71

<220>

<221> SITE

<222> (69)

<223> Xaa equals stop translation

<400> 110

Met Pro His Phe Leu Asp Trp Phe Val Pro Val Tyr Leu Val Ile Ser 1 5 10 15

Val Leu Ile Leu Val Gly Phe Gly Ala Cys Ile Tyr Tyr Phe Glu Pro 20 25 30

Gly Leu Gln Glu Ala His Lys Trp Arg Met Gln Arg Pro Leu Val Asp 35 40 45

Arg Xaa Leu Arg Lys Thr Leu Met Val Arg Asp Asn Leu Ala Phe Gly 50 55 60

Gly Pro Glu Val Xaa 65

<210> 111

<211> 123

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (123)

<223> Xaa equals stop translation

<400> 111

Met Ile Gly Gly Ile Thr Cys Ile Leu Ser Leu Ile Cys Ala Leu Ala 1 5 10 15

Leu Ala Tyr Leu Asp Gln Arg Ala Glu Arg Ile Leu His Lys Glu Gln
20 25 30

Gly Lys Thr Gly Glu Val Ile Lys Leu Thr Asp Val Lys Asp Phe Ser 35 40 45

Leu Pro Leu Trp Leu Ile Phe Ile Ile Cys Val Cys Tyr Tyr Val Ala 50 55 60

Val Phe Pro Phe Ile Gly Leu Gly Lys Val Phe Phe Thr Glu Lys Phe 65 70 75 80

Gly Phe Ser Ser Gln Ala Ala Ser Ala Ile Asn Ser Val Val Tyr Val
85 90 95

Ile Ser Ala Pro Met Ser Pro Val Phe Gly Leu Leu Val Asp Lys Thr 100 105 110

Gly Lys Asn Ile Ile Trp Val Leu Cys Ala Xaa 115 120

```
<210> 112
<211> 83
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (83)
<223> Xaa equals stop translation
<400> 112
Met Glu Lys Gln Cys Cys Ser His Pro Val Ile Cys Ser Leu Ser Thr
Met Tyr Thr Phe Leu Leu Gly Ala Ile Phe Ile Ala Leu Ser Ser Ser
Arg Ile Leu Leu Val Lys Tyr Ser Ala Asn Glu Gly Lys Leu Arg Leu
                              40
Gly Ile Cys Met Glu His Phe His Leu Ile Thr His Leu Ser Leu Ala
                         55
Phe Gly Ser Val Ile Tyr Asn Met Glu Ile Ile Met Pro Phe Ala Ser
Cys Glu Xaa
<210> 113
<211> 345
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (53)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (345)
<223> Xaa equals stop translation
<400> 113
Met Asp Phe Leu Val Leu Phe Leu Phe Tyr Leu Ala Ser Val Leu Met
                                      10
Gly Leu Val Leu Ile Cys Val Cys Ser Lys Thr His Ser Leu Lys Gly
                                  25
Leu Ala Arg Gly Gly Ala Gln Ile Phe Ser Cys Ile Ile Pro Glu Cys
Leu Gln Arg Ala Xaa His Gly Leu Leu His Tyr Leu Phe His Thr Arg
                                              60
                          55
```

Asn 65	His	Thr	Phe	Ile	Val 70	Leu	His	Leu	Val	Leu 75	Gln	Gly	Met	Val	Туr 80
Thr	Glu	Tyr	Thr	Trp 85	Glu	Val	Phe	Gly	Туг 90	Cys	Gln	Glu	Leu	Glu 95	Leu
Ser	Leu	His	туr 100	Leu	Leu	Leu	Pro	Туr 105	Leu	Leu	Leu	Gly	Val 110	Asn	Leu
Phe	Phe	Phe 115	Thr	Leu	Thr	Cys	Gly 120	Thr	Asn	Pro	Gly	11e 125	Ile	Thr	Lys
	130					135					140		Glu		
145					150					155			Lys		160
				165					170				Arg	175	
			180					185					Asn 190		
		195					200					205	Ala		
	210					215					220		Met		
225					230					235			His		240
				245					250				Pro	255	
			260					265					Leu 270		
		275					280					2,85	Thr		
	290					295					300		Pro		
305					310					315			Ile		320
His	Gly	Leu	Arg	Ser 325	Asn	Leu	Gln	Glu	11e 330		Leu	Pro	Ala	Phe 335	Pro
Cys	His	Glu	Arg 340	Lys	Lys	Gln	Glu	Xaa 345							

<210> 114 <211> 181

<212> PRT

74

<213> Homo sapiens <220> <221> SITE <222> (110) <223> Xaa equals any of the naturally occurring L-amino acids <400> 114 Met Ala Asp Pro His Val Ser Phe Leu Ser Phe Arg Gln Leu Phe Ser Trp Ala Ala Val Ile Leu Leu Arg Gly Ile Leu Gly Thr Val Ala Pro 25 Pro Pro Cys Pro Cys Val Leu Asp Leu Ala Val Tyr Pro Leu His Leu 35 40 Pro Val Glu Ala Pro Cys Leu Glu Val Val Phe Lys Gln Lys Asn Gly Lys Asp Asn Cys Leu Val Phe Tyr Pro Asp Pro Ile Pro Leu Arg Gly Ser Leu Leu Gly Pro Phe Ile Lys Asn Gln Cys His Ser Ser Val Ile Pro Leu Ser Asp Ser Ala Thr Ser Lys Ala Arg Ala Leu Xaa Leu Pro 105 Gly Arg Glu Thr Val Leu Ser Val Leu Pro Val Phe Ser Ser Pro Thr 115 120 Leu Pro Arg Thr His Ala Leu Gly Asp Ser Leu Gly Val Pro Gly Leu Leu Val Cys Ser Glu Thr Ser Thr Leu Asn Asp His Trp Cys Cys Arg 145 150 Arg Ala Gly Ala Tyr Ile Pro Ile Asn Arg Arg Phe Ser His Leu Met 170 Pro Leu Ala Phe Ser 180 <210> 115 <211> 116 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (116) <223> Xaa equals stop translation

Met Pro Ser Ser Ser Gly Leu Gly Ser Pro Ser Arg Pro Pro Ser

Ser Phe Leu Cys Leu Leu Leu Leu Leu Pro Pro Ala Ala Leu Ala 20 25 30

Leu Leu Phe Phe Leu Asp Phe Phe Pro Pro Arg Ala Ala Val Ser

Pro Phe Leu Pro Asp His Cys Ser Ala Arg Gln Pro Arg Val Trp Arg 50 55 60

Arg Glu Thr Leu Asn Arg Ser Ala Ser Gly Leu Gly Cys Trp Ala Arg 65 70 75 80

Ser Thr Glu Gln Gly Ala Val Gly Val Ala Thr Gly Thr Val Leu Asp
85 90 95

Ile Ser Leu Pro Ala Ser Cys Leu Ser Leu Trp Pro Pro Gly Pro Ser 100 105 110

Gly Gly Ile Xaa 115

<210> 116

<211> 71

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (71)

<223> Xaa equals stop translation

<400> 116

Met Asn Leu Leu Gly Met Ile Phe Ser Met Cys Gly Leu Met Leu Lys 1 5 10 15

Leu Lys Trp Cys Ala Trp Val Ala Val Tyr Cys Ser Phe Ile Ser Phe 20 25 30

Ala Asn Ser Arg Ser Ser Glu Asp Thr Lys Gln Met Met Ser Ser Phe 35 40 45

Met Leu Ser Ile Ser Ala Val Val Met Ser Tyr Leu Gln Asn Pro Gln 50 55 60

Pro Met Thr Pro Pro Trp Xaa 65 70

<210> 117

<211> 64

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (64)

WO 00/04140

76

<223> Xaa equals stop translation

<400> 117

Met Arg Asp Leu Ser Phe Leu Tyr Thr Leu Leu Trp Leu Pro Glu Ile

Trp Gln Ala Leu Ala Gly Gly Ile Arg Leu Asp Glu Val Glu Leu Leu

Glu Asn Glu Ala Val Leu Gly Glu Glu Met Arg Leu Tyr Arg Lys Ile

Asn Glu Val Val Leu Ser Gly Asn Glu Val Val Leu Gly Gly Lys Xaa 55

<210> 118

<211> 335

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (335)

<223> Xaa equals stop translation

<400> 118

Met Gly Ile Phe Pro Gly Ile Ile Leu Ile Phe Leu Arg Val Lys Phe 10

Ala Thr Ala Ala Val Ile Val Ser Gly Val Ser Lys His Leu His Cys

Ile Ser His Gln Lys Ser Thr Thr Val Ser His Glu Met Ser Gly Leu 40

Asn Trp Lys Pro Phe Val Tyr Gly Gly Leu Ala Ser Ile Val Ala Glu

Phe Gly Thr Phe Pro Val Asp Leu Thr Lys Thr Arg Leu Gln Val Gln

Gly Gln Ser Ile Asp Ala Arg Phe Lys Glu Ile Lys Tyr Arg Gly Met

Phe His Ala Leu Phe Arg Ile Cys Lys Glu Glu Gly Val Leu Ala Leu

Tyr Ser Gly Ile Ala Pro Ala Leu Leu Arg Gln Ala Ser Tyr Gly Thr 120

Ile Lys Ile Gly Ile Tyr Gln Ser Leu Lys Arg Leu Phe Val Glu Arg 130

Leu Glu Asp Glu Thr Leu Leu Ile Asn Met Ile Cys Gly Val Val Ser

145	•				150				-	155					160
Gly	Val	Ile	Ser	Ser 165	Thr	Ile	Ala	Asn	Pro 170	Thr	Asp	Val	Leu	Lys 175	Ile
Arg	Met	Gln	Ala 180	Gln	Gly	Ser	Leu	Phe 185	Gln	Gly	Ser	Met	Ile 190	Gly	Ser
Phe	Ile	Asp 195	Ile	Tyr	Gln	Gln	Glu 200	Gly	Thr	Arg	Gly	Leu 205	Trp	Arg	Glλ
Val	Val 210	Pro	Thr	Ala	Gln	Arg 215	Ala	Ala	Ile	Val	Val 220	Gly	Val	Glu	Leu
Pro 225	Val	Tyr	Asp	Ile	Thr 230	Lys	Lys	His	Leu	Ile 235	Leu	Ser	Gly	Met	Met 240
Gly	Asp	Thr	Ile	Leu 245	Thr	His	Phe	Val	Ser 250	Ser	Phe	Thr	Cys	Gly 255	Leu
Ala	Gly	Ala	Leu 260	Ala	Ser	Asn	Pro	Val 265	Asp	Val	Val	Arg	Thr 270	Arg	Met
Met	Asn	Gln 275	Arg	Ala	Ile	Val	Gly 280	His	Val	Asp	Leu	Tyr 285	Lys	Gly	Thr
Val	Asp 290	Gly	Ile	Leu	Lys	Met 295	Trp	Lys	His	Glu	Gly 300	Phe	Phe	Ala	Let
Tyr 305	Lys	Gly	Phe	Trp	Pro 310	Asn	Trp	Leu	Arg	Leu 315	Gly	Pro	Trp	Asn	11e 320
Ile	Phe	Phe	Ile	Thr 325	Tyr	Glu	Gln	Leu	Lys 330	Arg	Leu	Gln	Ile	Xaa 335	
<213 <213	0> 11 l> 22 2> PF B> Ho	21 RT	sapie	ens											
<222	l> s: 2> (5	51)	quals	s any	, of	the	nati	urall	ly o	ccuri	cing	L-ar	nino	acio	ds
	0> 11 Ala		Ala	Leu 5	Ala	Ala	Leu	Ala	Ala 10	Val	Glu	Pro	Ala	Cys 15	Gly
Ser	Arg	Tyr	Gln 20	Gln	Leu	Gln	Asn	Glu 25	Glu	Glu	Ser	Gly	Glu 30	Pro	Glu
Gln	Ala	Ala 35	Gly	Asp	Ala	Pro	Pro 40	Pro	Tyr	Ser	Ser	Ile 45	Ser	Ala	Glu
Ser	Ala 50	Xaa	Tyr	Phe	Asp	Tyr 55	Lys	Asp	Glu	Ser	Gly 60	Phe	Pro	Lys	Pro

78

Pro 65	Ser	Tyr	Asn	Val	Ala 70	Thr	Thr	Leu	Pro	Ser 75	Tyr	Asp	Glu	Ala	Glu 80
Arg	Thr	Lys	Ala	Glu 85	Ala	Thr	Ile	Pro	Leu 90	Val	Pro	Gly	Arg	Asp 95	Glu
Asp	Phe	Val	Gly 100	Arg	Asp	Asp	Phe	Asp 105	Asp	Ala	Asp	Gln	Leu 110	Arg	Ile
Gly	Asn	Asp 115	Gly	Ile	Phe	Met	Leu 120	Thr	Phe	Phe	Met	Ala 125	Phe	Leu	Phe
Asn	Trp 130	Ile	Gly	Phe	Phe	Leu 135	Ser	Phe	Суѕ	Leu	Thr 140	Thr	Ser	Ala	Ala
Gly 145	Arg	Туг	Gly	Ala	Ile 150	Ser	Gly	Phe	Gly	Leu 155	Ser	Leu	Ile	Lys	Trp 160
Ile	Leu	Ile	Val	Arg 165	Phe	Ser	Thr	Туг	Phe 170	Pro	Gly	Tyr	Phe	Asp 175	Gly
Gln	Туr	Trp	Leu 180	Trp	Trp	Val	Phe	Leu 185	Val	Leu	Gly	Phe	Leu 190	Leu	Phe
Leu	Arg	Gly 195		Ile	Asn	туг	Ala 200	Lys	Val	Arg	Lys	Met 205	Pro	Glu	Thr
Phe	Ser 210		Leu	Pro	Arg	Thr 215	Arg	Val	Leu	Phe	Ile 220	Tyr			
<21 <21 <21 <22	0>	73 PRT Iomo	sapi	ens											
<22		(473)		s st	op t	rans	alati	on							
<40 Met		120 s Phe	e Lev	ı Ile		e Ala	ı Phe	Phe	: Gly		/ Val	His	s Lev	Let 19	ı Ser
Leu	ı Cys	s Sei	20 20		s Ala	ı Ile	e Cys	ь Г уя	s Ası	n Gly	/ Ile	e Sei	r Lys 30	s Arg	y Thr
Phe	e Glu	u Gli		e Lys	s Glu	ı Glı	Ile د 40		a Se	r Cys	s Gly	Ası 4:	p Vai	l Ala	a Lys

Glu Arg Leu Ala Leu Leu Val Asp Thr Val Gly Pro Arg Leu Ser Gly
65 70 75 80

Ala Ile Ile Asn Leu Ala Val Tyr Gly Lys Ala Gln Asn Arg Ser Tyr

55

WO 00/04140

Ser	Lys	Asn	Leu	Glu 85	Lys	Ala	Ile	Gln	Ile 90	Met	Tyr	Gln	Asn	Leu 95	Glr
Gln	Asp	Gly	Leu 100	Glu	Lys	Va1	His	Leu 105	Glu	Pro	Val	Arg	Ile 110	Pro	His
Trp	Glu	Arg 115	Gly	Glu	Glu	Ser	Ala 120	Val	Met	Leu	Glu	Pro 125	Arg	Ile	His
Lys	Ile 130	Ala	Ile	Leu	Gly	Leu 135	Gly	Ser	Ser	Ile	Gly 140	Thr	Pro	Pro	Glu
Gly 145	Ile	Thr	Ala	Glu	Val 150	Leu	Val	Val	Thr	Ser 155	Phe	Asp	Glu	Leu	Glr 160
Arg	Arg	Ala	Ser	Glu 165	Ala	Arg	Gly	Lys	11e 170	Val	Val	Tyr	Asn	Gln 175	Pro
Tyr	Ile	Asn	Туг 180	Ser	Arg	Thr	Val	Gln 185	Tyr	Arg	Thr	Gln	Gly 190	Ala	Val
Glu	Ala	Ala 195	Lys	Val	Gly	Ala	Leu 200	Ala	Ser	Leu	Ile	Arg 205	Ser	Val	Ala
Ser	Phe 210	Ser	Ile	Туг	Ser	Pro 215	His	Thr	Gly	Ile	Gln 220	Glu	Туr	Gln	Asr
Gly 225	Val	Pro	Lys	Ile	Pro 230	Thr	Ala	Суѕ	Ile	Thr 235	Val	Glu	Asp	Ala	Glu 240
Met	Met	Ser	Arg	Met 245	Ala	Ser	His	Gly	11e 250	Lys	Ile	Val	Ile	Gln 255	Leu
Lys	Met	Gly	Ala 260	Lys	Thr	Tyr	Pro	Asp 265	Thr	Asp	Ser	Phe	Asn 270	Thr	Va]
Ala	Glu	11e 275	Thr	Gly	Ser	Lys	Туг 280	Pro	Glu	Gln	Val	Val 285	Leu	Val	Ser
Gly	His 290	Leu	Asp	Ser	Trp	Asp 295		Gly	Gln	Gly	Ala 300	Met	Asp	Asp	Glλ
Gly 305	Gly	Ala	Phe	Ile	Ser 310	Trp	Glu	Ala	Leu	Ser 315	Leu	Ile	Lys	Asp	120
Gly	Leu	Arg	Pro	Lys 325	Arg	Thr	Leu	Arg	Leu 330	Val	Leu	Trp	Thr	Ala 335	Glı
Glu	Gln	Gly	Gly 340	Val	Gly	Ala	Phe	Gln 345	Tyr	Tyr	Gln	Leu	His 350	Lys	Va]
Asn	Ile	Ser 355	Asn	Tyr	Ser	Leu	Val 360	Met	Glu	Ser	Asp	Ala 365	Gly	Thr	Phe
Leu	Pro 370	Thr	Gly	Leu	Gln	Phe 375	Thr	Gly	Ser	Glu	380	Ala	Arg	Ala	Ile
Mob	C1	C1	17-1	11 - h	C	t	1	C1 -	Dwo	T 011) cn	Tla	ጥኮሎ	Gln	Val

80 390 395 . 400 385 Leu Ser His Gly Glu Gly Thr Asp Ile Asn Phe Trp Ile Gln Ala Gly Val Pro Gly Ala Ser Leu Leu Asp Asp Leu Tyr Lys Tyr Phe Phe Phe 425 His His Ser His Gly Asp Thr Met Thr Val Met Asp Pro Lys Gln Met Asn Val Ala Ala Ala Val Trp Ala Val Val Ser Tyr Val Val Ala Asp 455 Met Glu Glu Met Leu Pro Arg Ser Xaa 470 <210> 121 <211> 168 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (168) <223> Xaa equals stop translation <400> 121 Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser 10 Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr 25 Pro Pro Ser Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu Gly Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro

Ile Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp 105 100

Asp Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu

Ala Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Ala 135

Gln Gln Arg Trp Lys Leu Gln Val Gln Glu Gln Arg Lys Ser Val Phe 145 150

. 81

Asp Arg His Val Val Leu Ser Xaa 165

<210> 122

<211> 47

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals stop translation

<400> 122

Met Lys Phe Ile Leu Trp Arg Arg Phe Arg Trp Ala Ile Ile Leu Phe 1 5 10 15

Ile Ile Leu Phe Ile Leu Leu Phe Leu Ala Ile Phe Ile Tyr Ala 20 25 30

Phe Pro Asn Tyr Ala Ala Met Lys Leu Val Lys Pro Phe Ser Xaa 35 40 45

<210> 123

<211> 108

<212> PRT

<213> Homo sapiens

<400> 123

Met His Gln Asp Trp Leu Cys Asn Leu Gly Trp Pro Leu Leu Ser Leu 1 5 10 15

Trp Ala Ala Glu Ser Ala Pro His Val Ala Met Ala Ser Ala Thr Ala 20 25 30

Gln Leu Trp Ser Arg Pro Cys Gly Arg Thr His Met Val Ser Leu Ala 35 40 45

Leu Gly His Gln Glu Thr Gly Leu Trp Leu Cys Ser Ala Phe Gly Cys 50 55 60

Val Val Asp Ser Pro Trp Ala Ser Val Cys Pro Ser Val Lys Gly Gln 65 70 75 80

Leu Thr Val Cys Gly Ile Leu Pro Arg Val Pro Val Cys Val Tyr Val
85 90 95

Cys Ala Cys Val Arg Val Ser Met Cys Val His Ile 100 105

<210> 124

<211> 60

<212> PRT

<213> Homo sapiens

WO 00/04140

82

<400> 124

Met Arg Gly Cys Val Pro Ala Phe Leu Leu His Val Leu Ser Leu Arg 10

Arg Ala Cys Cys Thr Gln Ala Ala Gln Val Phe Thr Ala Gln Leu Pro

Gly Arg Gln Val Ala Arg Arg Arg Gly Gly Trp His Glu Gln Gly 40

Gly Pro Met Leu Cys Ser Ser His His Ser Arg Thr 55 '

<210> 125

<211> 248

<212> PRT

<213> Homo sapiens

<400> 125

Met Ala Met Leu Pro Leu Val Leu His Trp Phe Phe Ile Glu Trp Tyr 10

Ser Gly Lys Lys Ser Ser Ser Ala Leu Phe Gln His Ile Thr Ala Leu 25

Phe Glu Cys Ser Met Ala Ala Ile Ile Thr Leu Leu Val Ser Asp Pro

Val Gly Val Leu Tyr Ile Arg Ser Cys Arg Val Leu Met Leu Ser Asp 55

Trp Tyr Thr Met Leu Tyr Asn Pro Ser Pro Asp Tyr Val Thr Thr Val 65

His Cys Thr His Glu Ala Val Tyr Pro Leu Tyr Thr Ile Val Phe Ile

Tyr Tyr Ala Phe Cys Leu Val Leu Met Met Leu Leu Arg Pro Leu Leu 105 100

Val Lys Lys Ile Ala Cys Gly Leu Gly Lys Ser Asp Arg Phe Lys Ser

Ile Tyr Ala Ala Leu Tyr Phe Phe Pro Ile Leu Thr Val Leu Gln Ala 135

Val Gly Gly Leu Leu Tyr Tyr Ala Phe Pro Tyr Ile Ile Leu Val 145 150

Leu Ser Leu Val Thr Leu Ala Val Tyr Met Ser Ala Ser Glu Ile Glu 170 165

Asn Cys Tyr Asp Leu Leu Val Arg Lys Lys Arg Leu Ile Val Leu Phe 180 185

Ser His Trp Leu Leu His Ala Tyr Gly Ile Ile Ser Ile Ser Arg Val

83

205

200

195

Asp Lys Leu Glu Gln Asp Leu Pro Pro Leu Ala Leu Val Pro Thr Pro 210 215 Ala Leu Phe Tyr Leu Phe Thr Ala Lys Phe Thr Glu Pro Ser Arg Ile 230 235 Leu Ser Glu Gly Ala Asn Gly His <210> 126 <211> 248 <212> PRT <213> Homo sapiens <400> 126 Met Glu Lys Ile Pro Glu Ile Gly Lys Phe Gly Glu Lys Ala Pro Pro Ala Pro Ser His Val Trp Arg Pro Ala Ala Leu Phe Leu Thr Leu Leu 25 Cys Leu Leu Leu Ile Gly Leu Gly Val Leu Ala Ser Met Phe His Val Thr Leu Lys Ile Glu Met Lys Lys Met Asn Lys Leu Gln Asn Ile Ser Glu Glu Leu Gln Arg Asn Ile Ser Leu Gln Leu Met Ser Asn Met 75 Asn Ile Ser Asn Lys Ile Arg Asn Leu Ser Thr Thr Leu Gln Thr Ile 85 Ala Thr Lys Leu Cys Arg Glu Leu Tyr Ser Lys Glu Gln Glu His Lys 105 Cys Lys Pro Cys Pro Arg Arg Trp Ile Trp His Lys Asp Ser Cys Tyr Phe Leu Ser Asp Asp Val Gln Thr Trp Gln Glu Ser Lys Met Ala Cys 135 Ala Ala Gln Asn Ala Ser Leu Leu Lys Ile Asn Asn Lys Asn Ala Leu 155 Glu Phe Ile Lys Ser Gln Ser Arg Ser Tyr Asp Tyr Trp Leu Gly Leu 165 Ser Pro Glu Glu Asp Ser Thr Arg Gly Met Arg Val Asp Asn Ile Ile 185 Asn Ser Ser Ala Trp Val Ile Arg Asn Ala Pro Asp Leu Asn Asn Met 195 Tyr Cys Gly Tyr Ile Asn Arg Leu Tyr Val Gln Tyr Tyr His Cys Thr

84 210 215 220 Tyr Lys Gln Arg Met Ile Cys Glu Lys Met Ala Asn Pro Val Gln Leu 235 Gly Ser Thr Tyr Phe Arg Glu Ala 245 <210> 127 <211> 612 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (245) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (246) -<223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (249) <223> Xaa equals any of the naturally occurring L-amino acids Met Ala Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro Leu Leu Ala Gly Leu Ala Leu Leu Gly Val Gly Pro Val Pro Ala Arg Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln 65 70 Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser 100 Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro

135

Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile

140

145					150					155					160
Met	Asp	Phe	Asn	Gly 165	Asp	Leu	Ile	Pro	Asp 170	Ile	Phe	Gly	Ile	Thr 175	Asn
Glu	Ser	Asn	Gln 180	Pro	Gln	Ile	Leu	Leu 185	Gly	Gly	Asn	Leu	Ser 190	Trp	His
Pro	Ala	Leu 195	Thr	Thr	Thr	Ser	Lys 200	Met	Arg	Ile	Pro	His 205	Ser	His	Ala
Phe	Ile 210	Asp	Leu	Thr	Glu	Asp 215	Phe	Thr	Ala	Asp	Leu 220	Phe	Leu	Thr	Thr
Leu 225	Asn	Ala	Thr	Thr	Ser 230	Thr	Phe	Gln	Phe	Glu 235	Ile	Trp	Glu	Asn	Leu 240
Asp	Gly	Asn	Phe	Xaa 245	Xaa	Ser	Thr	Xaa	Leu 250	Glu	Lys	Pro	Gln	Asn 255	Met
Met	Val	Val	Gly 260	Gln	Ser	Ala	Phe	Ala 265	Asp	Phe	Asp	Gly	Asp 270	Gly	His
Met	Asp	His 275	Leu	Leu	Pro	Gly	Cys 280	Glu	Asp	Lys	Asn	Cys 285	Gln	Lys	Ser
Thr	Ile 290	Tyr	Leu	Val	Arg	Ser 295	Gly	Met	Lys	Gln	Trp 300	Val	Pro	Val	Leu
Gln 305	Asp	Phe	Ser	Asn	Lys 310	Gly	Thr	Leu	Trp	Gly 315	Phe	Val	Pro	Phe	Val 320
Asp	Glu	Gln	Gln	Pro 325	Thr	Glu	Ile	Pro	Ile 330	Pro	Ile	Thr	Leu	His 335	Ile
Gly	Asp	Tyr	Asn 340	Met	Asp	Gly	Tyr	Pro 345	Asp	Ala	Leu	Val	11e 350	Leu	Lys
Asn	Thr	Ser 355	Gly	Ser	Asn	Gln	Gln 360	Ala	Phe	Leu	Leu	Glu 365	Asn	Val	Pro
Cys	Asn 370	Asn	Ala	Ser	Cys	Glu 375	Glu	Ala	Arg	Arg	Met 380	Phe	Lys	Val	Tyr
Trp 385	Glu	Leu	Thr	Asp	Leu 390	Asn	Gln	Ile	Lys	Asp 395	Ala	Met	Val	Ala	Thr 400
Phe	Phe	Asp	Ile	туr 405	Glu	Asp	Gly	Ile	Leu 410	Asp	Ile	Val	Val	Leu 415	Ser
Lys	Gly	Tyr	Thr 420	Lys	Asn	Asp	Phe	Ala 425	Ile	His	Thr	Leu	Lys 430	Asn	Asn
Phe	Glu	Ala 435	Asp	Ala	Tyr	Phe	Val 440	Lys	Val	Ile	Val	Leu 445	Ser	Gly	Leu
Cys	Ser 450	Asn	Asp	Cys	Pro	Arg 455	Lys	Ile	Thr	Pro	Phe 460	Gly	Val	Asn	Gln

Pro Gly Pro Tyr Ile Met Tyr Thr Thr Val Asp Ala Asn Gly Tyr Leu 475 470 Lys Asn Gly Ser Ala Gly Gln Leu Ser Gln Ser Ala His Leu Ala Leu 490 Gln Leu Pro Tyr Asn Val Leu Gly Leu Gly Arg Ser Ala Asn Phe Leu 505 500 Asp His Leu Tyr Val Gly Ile Pro Arg Pro Ser Gly Glu Lys Ser Ile Arg Lys Gln Glu Trp Thr Ala Ile Ile Pro Asn Ser Gln Leu Ile Val 530 Ile Pro Tyr Pro His Asn Val Pro Arg Ser Trp Ser Ala Lys Leu Tyr 555 550 Leu Thr Pro Ser Asn Ile Val Leu Leu Thr Ala Ile Ala Leu Ile Gly 565 Val Cys Val Phe Ile Leu Ala Ile Ile Gly Ile Leu His Trp Gln Glu 585 Lys Lys Ala Asp Asp Arg Glu Lys Arg Gln Glu Ala His Arg Phe His Phe Asp Ala Met 610 <210> 128 <211> 447 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (8) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (28) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (309) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (333) <223> Xaa equals any of the naturally occurring L-amino acids <400> 128

Met 1	Glu	Leu	Ser	Gln 5	Met	Ser	Xaa	Leu	Met 10	Gly	Leu	Ser	Val	Leu 15	Lev
Gly	Leu	Leu	Ala 20	Leu	Met	Ala	Thr	Ala 25	Ala	Val	Xaa	Arg	Gly 30	.Trp	Le
Arg	Ala	Gly 35	Glu	Glu	Arg	Ser	Gly 40	Arg	Pro	Ala	Cys	Gln 45	Lys	Ala	Ası
Gly	Phe 50	Pro	Pro	Asp	Lys	Ser 55	Ser	Gly	Ser	Lys	Lys 60	Gln	Lys	Gln	Туі
Gln 65	Arg	Ile	Arg	Lys	Glu 70	Lys	Pro	Gln	Gln	His 75	Asn	Phe	Thr	His	Arg 80
Leu	Leu	Ala	Ala	Ala 85	Leu	Lys	Ser	His	Ser 90	Gly	Asn	Ile	Ser	Суs 95	Met
Asp	Phe	Ser	Ser 100	Asn	Gly	Lys	Tyr	Leu 105	Ala	Thr	Суѕ	Ala	Asp 110	Asp	Aŗg
Thr	Ile	Arg 115	Ile	Trp	Ser	Thr	Lys 120	Asp	Phe	Leu	Gln	Arg 125	Glu	His	Arg
Ser	Met 130	Arg	Ala	Asn	Val	Glu 135	Leu	Asp	His	Ala	Thr 140	Leu	Val	Arg	Phe
Ser 145	Pro	Asp	Cys	Arg	Ala 150	Phe	Ile	Val	Trp	Leu 155	Ala	Asn	Gly	Asp	Thr 160
Leu	Arg	Val	Phe	Lys 165	Met	Thr	Lys	Arg	Glu 170	Asp	Gly	Gly	Tyr	Thr 175	Phe
Thr	Ala	Thr	Pro 180	Glu	Asp	Phe	Pro	Lys 185	Lys	His	Lys	Ala	Pro 190	Val	Ile
Asp	Ile	Gly 195	Ile	Ala	Asn	Thr	Gly 200	Lys	Phe	Ile	Met	Thr 205	Ala	Ser	Ser
Asp	Thr 210	Thr	Val	Leu	Ile	Trp 215	Ser	Leu	Lys	Gly	Gln 220	Val	Leu	Ser	Thr
Ile 225	Asn	Thr	Asn	Gln	Met 230	Asn	Asn	Thr	His	Ala 235	Ala	Val	Ser	Pro	Cys 240
Gly	Arg	Phe	Val	Ala 245	Ser	Cys	Gly	Phe	Thr 250	Pro	Asp	Val	Lys	Val 255	Trp
Glu	Val	Cys	Phe 260	Gly	Lys	Lys	Gly	Glu 265	Phe	Gln	Glu	Val	Val 270	Arg	Ala
Phe	Glu	Leu 275	Lys	Gly	His	Ser	Ala 280	Ala	Val	His	Ser	Phe 285	Ala	Phe	Ser
Asn	Asp 290	Ser	Arg	Arg	Met	Ala 295	Ser	Val	Ser	Lys	Asp 300	Gly	Thr	Trp	Lys
Leu	Trp	Asp	Thr	Xaa	Val	Glu	Tyr	Lys	Lys	Lys	Gln	Asp	Pro	Tyr	Leu

315 320 305 310 Leu Lys Thr Gly Arg Phe Glu Glu Ala Ala Gly Ala Xaa Pro Cys Arg 330 325 Leu Ala Leu Ser Pro Asn Ala Gln Val Leu Ala Leu Ala Ser Gly Ser 345 340 Ser Ile His Leu Tyr Asn Thr Arg Arg Gly Glu Lys Glu Glu Cys Phe 360 Glu Arg Val His Gly Glu Cys Ile Ala Asn Leu Ser Phe Asp Ile Thr Gly Arg Phe Leu Ala Ser Cys Gly Asp Arg Ala Val Arg Leu Phe His 395 390 Asn Thr Pro Gly His Arg Ala Met Val Glu Glu Met Gln Gly His Leu Lys Arg Ala Ser Asn Glu Ser Thr Arg Gln Arg Leu Gln Gln Gln Leu Thr Gln Ala Gln Glu Thr Leu Lys Ser Leu Gly Ala Leu Lys Lys 435 440 <210> 129 <211> 291 <212> PRT <213> Homo sapiens <400> 129 Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg 40 Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys Pro Trp Val Gly Asn Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe 120 115 Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu

<212> PRT

89 140 130 135 Thr Leu Lys Glu Thr Pro Gly Thr Val Leu Glu Val Leu Ile Cys Phe 150 155 Phe Thr Leu Trp Ser Val Val Gly Leu Thr Gly Phe His Thr Phe Leu Val Ala Leu Asn Gln Thr Thr Asn Glu Asp Ile Lys Gly Ser Trp Thr 185 Gly Lys Asn Arg Val Gln Asn Pro Tyr Ser His Gly Asn Ile Val Lys 195 200 Asn Cys Cys Glu Val Leu Cys Gly Pro Leu Pro Pro Ser Val Leu Asp 215 Arg Arg Gly Ile Leu Pro Leu Glu Glu Ser Gly Ser Arg Pro Pro Ser 230 Thr Gln Glu Thr Ser Ser Ser Leu Leu Pro Gln Ser Pro Ala Pro Thr 250 Glu His Leu Asn Ser Asn Glu Met Pro Glu Asp Ser Ser Thr Pro Glu Glu Met Pro Pro Glu Pro Pro Glu Pro Pro Gln Glu Ala Ala Glu 280 275 Ala Glu Lys 290 <210> 130 <211> 78 <212> PRT <213> Homo sapiens <400> 130 Met Val Arg Lys Trp Leu Thr Phe Val Glu His Leu Leu Cys Ala Trp Pro Arg Leu Gly Ala Phe Val Pro Arg Val Thr Pro Ser Glu Cys Ser . 20 Ser Leu Pro His Ser Asn Trp Gly Val Gly Gly Arg Ala Ala Gln Leu 40 Thr Gly Ala Glu Leu Lys Thr His Ser Trp Val Cys Leu Gly Trp Ala Val Leu Val Ala Pro Val Ala Asn Thr Arg Ala Pro Phe Thr 70 <210> 131 <211> 333

PCT/US99/15849 WO 00/04140

90

<213> Homo sapiens

<220>

<221> SITE

<222> (97)

<223> Xaa equals any of the naturally occurring L-amino acids

Met Leu Met Phe Ala Val Ile Val Ala Ser Ser Gly Leu Leu Met 10

Ile Glu Arg Gly Ile Leu Ala Glu Met Lys Pro Leu Pro Leu His Pro

Pro Gly Arg Glu Gly Thr Ala Trp Arg Gly Lys Ala Pro Lys Pro Gly

Gly Leu Ser Leu Arg Ala Gly Asp Ala Asp Leu Gln Val Arg Gln Asp

Val Arg Asn Arg Thr Leu Arg Ala Val Cys Gly Gln Pro Gly Met Pro

Arg Asp Pro Trp Asp Leu Pro Val Gly Gln Arg Arg Thr Leu Leu Arg 85

Xaa Ile Leu Val Ser Asp Arg Tyr Arg Phe Leu Tyr Cys Tyr Val Pro 105

Lys Val Ala Cys Ser Asn Trp Lys Arg Val Met Lys Val Leu Ala Gly

Val Leu Asp Ser Val Asp Val Arg Leu Lys Met Asp His Arg Ser Asp 135

Leu Val Phe Leu Ala Asp Leu Arg Pro Glu Glu Ile Arg Tyr Arg Leu 155 150

Gln His Tyr Phe Lys Phe Leu Phe Val Arg Glu Pro Leu Glu Arg Leu 170 165

Leu Ser Ala Tyr Arg Asn Lys Phe Gly Glu Ile Arg Glu Tyr Gln Gln 185

Arg Tyr Gly Ala Glu Ile Val Arg Arg Tyr Arg Ala Gly Ala Gly Pro 195

Ser Pro Ala Gly Asp Asp Val Thr Phe Pro Glu Phe Leu Arg Tyr Leu 215

Val Asp Glu Asp Pro Glu Arg Met Asn Glu His Trp Met Pro Val Tyr 225 230

His Leu Cys Gln Pro Cys Ala Val His Tyr Asp Phe Val Gly Ser Tyr 250 245

Glu Arg Leu Glu Ala Asp Ala Asn Gln Val Leu Glu Trp Val Arg Ala 265

Pro Pro His Val Arg Phe Pro Ala Arg Gln Ala Trp Tyr Arg Pro Ala 275 280 285

Ser Pro Glu Ser Leu His Tyr His Leu Cys Ser Ala Pro Arg Ala Leu 290 295 300

Leu Gln Asp Val Leu Pro Lys Tyr Ile Leu Asp Phe Ser Leu Phe Ala 305 310 315 320

Tyr Pro Leu Pro Asn Val Thr Lys Glu Ala Cys Gln Gln 325 330

<210> 132

<211> 164

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (126)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 132

Met Leu Pro Leu Leu Ile Ile Cys Leu Leu Pro Ala Ile Glu Gly Lys
1 5 10 15

Asn Cys Leu Arg Cys Trp Pro Glu Leu Ser Ala Leu Ile Asp Tyr Asp 20 25 30

Leu Gln Ile Leu Trp Val Thr Pro Gly Pro Pro Thr Glu Leu Ser Gln 35 40 45

Ser Ile His Ser Leu Phe Leu Glu Asp Asn Asn Phe Leu Lys Pro Trp

Tyr Leu Asp Arg Asp His Leu Glu Glu Glu Thr Ala Lys Phe Phe Thr
65 70 75 80

Gln Val His Gln Ala Ile Lys Thr Leu Arg Asp Asp Lys Thr Val Leu 85 90 95

Leu Glu Glu Ile Tyr Thr His Lys Asn Leu Phe Thr Glu Arg Leu Asn 100 105 110

Lys Ile Ser Asp Gly Leu Lys Glu Lys Gly Ala Pro Pro Xaa Ser Met 115 120 125

Asn Ala Phe Pro Ala Pro Ser Pro Thr Cys Thr Pro Glu Pro Leu Gly
130 135 140

Ser Val Cys Leu Pro Ser Thr Ser Val Ser Leu Pro Ser His Leu Pro 145 150 155 160

Gly Ser Leu Gln

<210: <211: <212: <213	> 24 > PR	5 T	apie	ns											
<220 <221 <222 <223	> SI > (2	45)	uals	sto	p tr	ansl	atio	n							
<400 Met 1	> 13 Val	3 Ala	Val	Gly 5	Val	Tyr	Ala	Arg _.	Leu 10	Met	Lys	His	Ala	Glu 15	Ala
Ala	Leu	Ala	Cys 20	Leu	Ala	Val	Asp	Pro 25	Ala	Ile	Leu	Leu	11e 30	Val	Val
Gly	Val	Leu 35	Met	Phe	Leu	Leu	Thr 40	Phe	Cys	Gly	Суѕ	Ile 45	Gly	Ser	Leu
Arg	Glu 50	Asn	Ile	Cys	Leu	Leu 55	Gln	Thr	Phe	Ser	Leu 60	Cys	Leu	Thr	Ala
Val 65	Phe	Leu	Leu	Gln	Leu 70	Ala	Ala	Gly	Ile	Leu 75	Gly	Phe	Val	Phe	Ser 80
Asp	Lys	Ala	Arg	Gly 85	Lys	Val	Ser	Glu	Ile 90	Ile	Asn	Asn	Ala	Ile 95	Val
His	Tyr	Arg	Asp 100	Asp	Leu	Asp	Leu	Gln 105	Asn	Leu	Ile	Asp	Phe 110	Gly	Gln
Lys	Lys	Phe 115	Ser	Cys	Cys	Gly	Gly 120	Ile	Ser	Туг	Lys	Asp 125	Trp	Ser	Gln
Asn	Met 130		Phe	Asn	Cys	Ser 135	Glu	Asp	Asn	Pro	Ser 140	Arg	Glu	Arg	Cys
Ser 145		Pro	Туr	Ser	Cys 150	Cys	Leu	Pro	Thr	Pro	Asp	Gln	Ala	Val	Ile 160
Asn	Thr	Met	Cys	Gly 165	Gln	Gly	Met	Gln	Ala 170	Phe	. Asp	Tyr	Leu	. Glu 175	Ala
Ser	Lys	. Val	. Ile		Thr	Asn	Gly	7 Cys	i Ile	e Asp	Lys	Lev	Val 190	. Asn	Trp
Ile	e His	Ser 195		Lev	ı Phe	e Lev	Let 200	ı Gly	/ Gly	/ Val	L Ala	Let 205	i Gly	, Leu	Ala
Ile	e Pro	Glr	ı Lev	ı Val	l Gly	/ Ile	e Lev	ı Lev	ı Sei	r Glr	n Ile 220	e Lei	ı Val	. Asr	Gln

Ile Lys Asp Gln Ile Lys Leu Gln Leu Tyr Asn Gln Gln His Arg Ala

230

235

Asp Pro Trp Tyr Xaa

```
<210> 134
<211> 56
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation
<400> 134
Met Gly Thr Val Gly Leu Trp Pro Ser Trp Leu Trp Leu Pro Ala Ser
                                    10
Trp Pro Leu Thr Ser Cys Gly Val Thr Arg Arg Arg Leu Arg Gly Pro
Gly Leu Arg Arg Thr Ser Gln Thr Gly Arg His Thr Ser Pro Cys Pro
                             40
Thr Ala Thr Trp Ala Glu Ser Xaa
     50
<210> 135
<211> 55
<212> PRT
<213> Homo sapiens
<220> 
<221> SITE
<222> (47)
<223> Xaa equals any of the naturally occurring L-amino acids
 <220>
 <221> SITE
 <222> (51)
 <223> Xaa equals any of the naturally occurring L-amino acids
 <220>
 <221> SITE
 <222> (55)
 <223> Xaa equals stop translation
 <400> 135
 Met Ser Ile Val Met Ser Pro Leu Leu Pro Ile Cys Tyr Leu Asn
   1
                   5
 Leu Leu Phe Phe Val Asn Leu Ala Lys Asn Leu Ser Ile Leu Phe
                                  25
 Val Ser Ser Lys Lys Tyr Thr Phe Val Phe Met Ile Ser Leu Xaa Phe
                                                   45
                              40
          35 '
```

Phe His Xaa Tyr Phe Ile Xaa

94

50 55

<210> 136

<211> 89

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (89)

<223> Xaa equals stop translation

<400> 136

Met Ala Ile Ile Ser Phe Glu Leu Leu Phe Leu Met Asn Leu Pro Thr 1 5 10 15

Val Asn Ser Ser Asn Phe Lys Leu Ile Ile Pro Glu Asp Val Thr Leu 20 25 30

Ser Phe Val Ser His Leu Asp Ile Thr Val Asn His Phe Val Phe Leu 35 40 45

Ser Thr Phe Glu Leu Ala Gly Val Ile Glu Gly Lys Pro Leu Pro Asp 50 55 60

Ser Lys Ser Asp Leu Cys Pro Ile Leu Gly Gln Leu Trp Phe His Ile 65 70 75 80

Leu Leu Phe Phe Ile Phe Trp Val Xaa 85

<210> 137

<211> 62

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (62)

<223> Xaa equals stop translation

<400> 137

Met Arg Leu Pro Ile Ala Pro His Leu Gln Tyr Phe Met Trp Ser Val 1 5 10 15

Leu Leu Phe Leu Val Ile Leu Val Asp Met Lys Trp His Leu Ser Val 20 25 30

Ala Phe His Tyr Ile Ser Leu Met Thr Asn Gly Ile Leu Ser Pro Phe 35 40 45

Gln Cys Leu Leu Ala Ile His Val Ser Leu Phe Phe Val Xaa 50 60

95

<211> 106

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (106)

<223> Xaa equals stop translation

<400> 138

Met Cys Leu Leu Pro Gly Gly Val Leu Leu Ile Trp Ser Cys Ala Ser 1 5 10 15

Gly Thr Pro Ala Ser His Thr Lys Asp Trp Gly Arg Cys Lys Phe Ser

Ala Ala Thr Lys Arg Thr Ala Glu Ser Asn Leu Glu Ser Thr Gln Leu
35 40 45

Met Leu Ala Ser Gln Ile Asp Pro Leu Leu Ala Glu Cys Trp His Leu 50 55 60

Cys Ala Ser Val Ser Ser Val Asn Gly Gly Asp Lys Lys Cys Val 65 70 75 80

His Thr Ser Arg Ala Val Gly Arg Ile Lys Leu Cys Ser Asp Thr Ile 85 90 95

Arg Ala Cys Ser Gly Trp Tyr Leu Gln Xaa 100 105

<210> 139

<211> 52

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (52)

<223> Xaa equals stop translation

<400> 139

Met Ser His Ser Val Phe Ala His Tyr Ile Phe Asn Ile Leu Leu Leu 1 5 10 15

Leu Leu Leu Leu Leu Ile Gly Phe Leu Tyr Ser Met Pro Phe Ile
20 25 30

Tyr Lys Asp Thr Lys Lys Thr His Val Cys Asn Phe Asn Asn Ile Phe 35 40 45

Pro Ile Leu Xaa 50

<210> 140

<211> 119

.....

96

<212> PRT

<213> Homo sapiens

<400> 140

Met Lys Trp Arg Arg Lys Ser Ala Tyr Trp Lys Ala Leu Lys Val Phe 1 5 10 15

Lys Leu Pro Val Glu Phe Leu Leu Leu Leu Thr Val Pro Val Val Asp 20 25 30

Pro Asp Lys Asp Gln Asn Trp Lys Arg Pro Leu Asn Cys Leu His
35 40 45

Leu Val Ile Ser Pro Leu Val Val Val Leu Thr Leu Gln Ser Gly Thr
50 60

Tyr Gly Val Tyr Glu Ile Gly Gly Leu Val Pro Val Trp Val Val Val 65 70 75 80

Val Ile Ala Gly Thr Ala Leu Ala Ser Val Thr Phe Phe Ala Thr Ser 85 90 95

Asp Ser Gln Pro Pro Arg Leu His Trp Leu Phe Ala Phe Leu Gly Phe
100 105 110

Leu Thr Ser Ala Leu Trp Ile 115

<210> 141

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals stop translation

<400> 141

Met Cys Ser Gly Ser Phe Lys Glu Leu Tyr Leu Val Pro Ile Ser Leu 1 5 10 15

Phe Ser Thr Cys Val Leu Gly Phe Tyr Phe His Asn Phe Leu Leu Leu 20 25 30

Ile Ile Leu Phe Ser Ile Leu Leu Arg Lys Ile Thr Gly Lys Leu Phe 35 40 45

Phe Thr Tyr Tyr His Phe Ser Cys Gly Val Xaa 50 55

<210> 142

<211> 100

<212> PRT

<213> Homo sapiens

97

```
<220>
<221> SITE
<222> (100)
<223> Xaa equals stop translation
<400> 142
Met Leu Phe Phe Leu Ser Leu Phe Leu
```

Met Leu Phe Phe Leu Ser Leu Phe Leu Ser Leu Leu Leu Thr Leu Ser 1 5 10 15

Leu Pro Ser Phe Leu Pro Phe Ser Phe Phe Phe Ser Leu Phe Pro 20 25 30

His Leu Ser Ala Cys Leu Leu Pro Ser Leu Pro Ser Pro Pro Phe Pro 35 40 45

Leu Pro Pro Ser Leu Pro Ser Phe Leu Pro Ser Phe 50 55 60

Leu Pro Ser Leu Leu Ser Pro Ser Phe Pro Ala Phe Pro Ser Phe 65 70 75 80

Cys Gln Leu Ala Arg Arg Ser Pro Arg Lys Ser Thr Gln Met Leu Gln 85 90 95

Ser Thr Ser Xaa 100

<210> 143 <211> 65 <212> PRT

<213> Homo sapiens

<220>
<221> SITE
<222> (61)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (65)
<223> Xaa equals stop translation

<400> 143
Met Ala Val Leu Leu Ile Thr Ile Leu Leu Phe Leu Cys Leu Gly Tyr
1 5 10 15

Tyr Arg Val Ile Thr Glu Ile Ser Arg Lys Thr Pro Ala Cys Arg Met 20 25 30

Phe Thr Ser Ser Leu Ser Ser Trp Tyr Ile Met Arg Lys Leu Tyr Asp 35 40 45

Thr Pro Gly Glu Val Phe Leu Ser His Ala Ile Val Xaa Phe Leu Lys 50 55 60

Xaa 65

```
<210> 144
<211> 67
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (67)
<223> Xaa equals stop translation
<400> 144
Met Leu Asn Gln Pro Cys Ile Leu Gly Met Lys Pro Thr Trp Leu Trp
Trp Ile Ser Phe Leu Met Cys Cys Trp Val Trp Leu Ala Ser Val Leu
Leu Gly Ile Phe Ala Ser Ile Phe Ile Arg Asp Ile Gly Leu Glu Phe
                                                  45
         35
                              40
Ser Phe Phe Val Met Cys Leu Pro Gly Phe Gly Ile Arg Val Met Leu
Ala Ser Xaa
 65
<210> 145
<211> 59
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (59)
<223> Xaa equals stop translation
<400> 145
Met Thr Ala Met Ser Ile His Leu Phe Cys Thr Ala Leu Ser Cys Gly
Ser Ser Gly Gln Cys Asn Lys Ala Ile Lys Arg Asn Lys Ile Ser Asn
Asp Trp Lys Asp Val Asn Val Ser Ser Phe Ile Glu Asn Met Ile His
         35
Arg Tyr Thr Tyr Thr Asn Ala Leu Asn Ser Xaa
     50
                         55
<210> 146
<211> 56
<212> PRT
<213> Homo sapiens
```

the second section of the section of the second section of the second section of the section of the second section of the section of th

99

```
<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation
<400> 146
Met Ser His Cys Thr Trp Pro Val Cys Leu Phe Cys Leu Val Pro Pro
Pro Met Gly Asp Leu Lys Glu Val Cys Leu Pro His Arg Cys Pro Gly
                                  25
Arg Thr Ala Cys Cys Ser Tyr Ser Glu Pro His Leu Gln Thr Glu Glu
                             40
Asp Arg Arg Thr Leu Ile Cys Xaa
<210> 147
<211> 66
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (66)
<223> Xaa equals stop translation
Met Thr Asn Gly His Gln Val Leu Leu Leu Leu Leu Leu Thr Ser Ala
Val Ala Ala Gly Pro Trp Pro Gln Val His Ala Gly Gln Trp Gly Trp
Met Cys Leu Pro Pro Gly Leu Pro Ser Val Gln Ala Arg Ser Gly Leu
Gly Gly Leu Pro Gly Gly Pro Gln Trp Val Pro Gly Gly Ala Arg Gly
Tyr Xaa
 65
<210> 148
<211> 328
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (328)
<223> Xaa equals stop translation
```

Met Ala Cys Arg Lys Leu Ala Val Ala His Pro Leu Leu Leu Arg

1,00

1				5					10					15	
His	Leu	Pro	Met 20	Ile	Ala	Ala	Leu	Leu 25	His	Gly	Arg	Thr	His 30	Leu	Asn
Phe	Gln	Glu 35	Phe	Arg	Gln	Gln	Asn 40	His	Leu	Ser	Суѕ	Phe 45	Leu	His	Val
Leu	Gly 50	Leu	Leu	Glu	Leu	Leu 55	Gln	Pro	His	Val	Phe 60	Arg	Ser	Glu	His
Gln 65	Gly	Ala	Leu	Trp	Asp 70	Суѕ	Leu	Leu	Ser	Phe 75	Ile	Arg	Leu	Leu	Leu 80
Asn	Tyr	Arg	Lys	Ser 85	Ser	Arg	His	Leu	Ala 90	Ala	Phe	Ile	Asn	Lys 95	Phe
Val	Gln	Phe	Ile 100	His	Lys	Tyr	Ile	Thr 105	Tyr	Asn	Ala	Pro	Ala 110	Ala	Ile
Ser	Phe	Leu 115	Gln	Lys	His	Ala	Asp 120	Pro	Leu	His	Asp	Leu 125	Ser	Phe	Asp
Asn	Ser 130	Asp	Leu	Val	Met	Leu 135	Lys	Ser	Leu	Leu	Ala 140	Gly	Leu	Ser	Leu
Pro 145	Ser	Arg	Asp	Asp	Arg 150	Thr	Asp	Arg	Gly	Leu 155	Asp	Glu	Glu	Gly	Glu 160
Glu	Glu	Ser	Ser	Ala 165	Gly	Ser	Leu	Pro	Leu 170	Val	Ser	Val	Ser	Leu 175	Phe
Thr	Pro	Leu	Thr 180	Ala	Ala	Glu	Met	Ala 185	Pro	Tyr	Met	Lys	Arg 190	Leu	Ser
Arg	Gly	Gln 195	Thr	Val	Glu	Asp	Leu 200	Leu	Glu	Val	Leu	Ser 205	Asp	Ile	Asp
Glu	Met 210	Ser	Arg	Arg	Arg	Pro 215	Glu	Ile	Leu	Ser	Phe 220	Phe	Ser	Thr	Asn
Leu 225	Gln	Arg	Leu	Met	Ser 230	Ser	Ala	Glu	Glu	Cys 235	Cys	Arg	Asn	Leu	Ala 240
Phe	Ser	Leu	Ala	Leu 245	Arg	Ser	Met	Gln	Asn 250	Ser	Pro	Ser	Ile	Ala 255	Ala
Ala	Phe	Leu	Pro 260	Thr	Phe	Met	Tyr	Cys 265	Leu	Gly	Ser	Gln	Asp 270	Phe	Glu
Val	Val	Gln 275	Thr	Ala	Leu	Arg	Asn 280	Leu	Pro	Glu	Tyr	Ala 285	Leu	Leu	Суѕ
Gln	Glu 290	His	Ala	Ala	Val	Leu 295	Leu	His	Arg	Ala	Phe 300	Leu	Val	Gly	Met
Tyr 305	Gly	Gln	Met	Asp	Pro 310	Ser	Ala	Gln	Ile	Ser 315	Glu	Ala	Leu	Arg	11e 320

Leu His Met Glu Ala Val Met Xaa 325

<222> (149)

<400> 150

<223> Xaa equals stop translation

5

```
<210> 149
<211> 90
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (10)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (90)
<223> Xaa equals stop translation
Met Gly Phe Leu Gln Leu Leu Val Val Xaa Val Leu Xaa Ser Glu His
Arg Val Ala Gly Ala Ala Glu Val Phe Gly Asn Ser Ser Glu Gly Leu
Ile Glu Phe Ser Val Gly Lys Phe Arg Tyr Phe Glu Leu Asn Arg Pro
Phe Pro Glu Glu Ala Ile Leu His Asp Ile Ser Ser Asn Val Thr Phe
                         55
Leu Ile Phe Gln Ile His Ser Gln Tyr Gln Asn Thr Thr Val Ser Phe
                                         75
                     70
Ser Pro Arg Arg Ser Pro Thr Met Xaa
<210> 150
<211> 149
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
```

Met Ala Gly Ser Pro Leu Leu Trp Gly Pro Arg Ala Gly Gly Val Gly

Leu Leu Val Leu Leu Leu Gly Leu Phe Arg Pro Pro Pro Ala Leu 20 25 30

Cys Ala Arg Pro Val Lys Glu Pro Arg Gly Leu Ser Ala Ala Ser Pro 35 40 45

Pro Leu Ala Arg Leu Ala Leu Leu Ala Ala Ser Gly Gln Cys Pro 50 60

Glu Val Arg Arg Gly Arg Cys Arg Pro Gly Ala Gly Ala Gly Ala 65 70 75 80

Ser Ala Gly Ala Glu Arg Gln Glu Arg Ala Arg Ala Glu Ala Gln Arg 85 90 95

Leu Arg Ile Ser Arg Arg Ala Ser Trp Arg Ser Cys Cys Ala Ser Gly
100 105 110

Ala Pro Pro Ala Thr Leu Ile Arg Leu Trp Ala Trp Thr Thr Pro 115 120 125

Thr Arg Leu Gln Arg Ser Ser Leu Ala Leu Cys Ser Ala Pro Ala Leu 130 135 140

Thr Leu Pro Pro Xaa 145

<210> 151

<211> 391

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (391)

<223> Xaa equals stop translation

<400> 151

Met Leu Pro Thr Trp Trp Ile Val Ser Ser Trp Leu Val Trp Gly Val
1 5 10 15

Ile Leu Phe Val Tyr Leu Val Ile Arg Ala Leu Arg Leu Trp Arg Thr 20 25 30

Ala Lys Leu Gln Val Thr Leu Lys Lys Tyr Ser Val His Leu Glu Asp 35 40 45

Met Ala Thr Asn Ser Arg Ala Phe Thr Asn Leu Val Arg Lys Ala Leu 50 55 60

Arg Leu Ile Gln Glu Thr Glu Val Ile Ser Arg Gly Phe Thr Leu Val 65 70 75 80

Ser Ala Ala Cys Pro Phe Asn Lys Ala Gly Gln His Pro Ser Gln His 85 90 95

Leu	Ile	Gly	Leu 100	Arg	Lys	Ala	Val	Туг 105	Arg	Thr	Leu	Arg	Ala 110	Asn	Phe
Gln	Ala	Ala 115	Arg	Leu	Ala	Thr	Leu 120	Tyr	Met	Leu	Lys	Asn 125	Tyr	Pro	Leu
Asn	Ser 130	Glu	Ser	Asp	Asn	Val 135	Thr	Asn	Tyr	Ile	Cys 140	Val	Val	Pro	Phe
Lys 145	Glu	Leu	Gly	Leu	Gly 150	Leu	Ser	Glu	Glu	Gln 155	Ile	Ser	Glu	Glu	Glu 160
Ala	His	Asn	Phe	Thr 165	Asp	Gly	Phe	Ser	Leu 170	Pro	Ala	Leu	Lys	Val 175	Leu
Phe	Gln	Leu	Trp 180	Val	Ala	Gln	Ser	Ser 185	Glu	Phe	Phe	Arg	Arg 190	Leu	Ala
Leu	Leu	Leu 195	Ser	Thr	Ala	Asn	Ser 200	Pro	Pro	Gly	Pro	Leu 205	Leu	Thr	Pro
Ala	Leu 210	Leu	Pro	His	Arg	11e 215	Leu	Ser	Asp	Val	Thr 220	Gln	Gly	Leu	Pro
His 225	Ala	His	Ser	Ala	Cys 230	Leu	Glu	Glu	Leu	Lys 235	Arg	Ser	Tyr	Glu	Phe 240
Tyr	Arg	Tyr	Phe	Glu 245	Thr	Gln	His	Gln	Ser 250	Val	Pro	Gln	Cys	Leu 255	Ser
Lys	Thr	Gln	Gln 260	Lys	Ser	Arg	Glu	Leu 265	Asn	Asn	Val	His	Thr 270	Ala	Val
Arg	Ser	Leu 275	Gln	Leu	His	Leu	Lys 280	Ala	Leu	Leu	Asn	Glu 285	Val	Ile	Ile
Leu	Glu 290	Asp	Glu	Leu	Glu	Lys 295	Leu	Val	Cys	Thr	Lys 300	Glu	Thr	Gln	Glu
Leu 305	Val	Ser	Glu	Ala	Туг 310	Pro	Ile	Leu	Glu	Gln 315	Lys	Leu	Lys	Leu	11e 320
Gln	Pro	His	Val	Gln 325	Ala	Ser	Asn	Asn	Суs 330	Trp	Glu	Glu	Ala	Ile 335	Ser
Gln	Val	Asp	Lys 340	Leu	Leu	Arg	Arg	Asn 345	Thr	Asp	Lys	Lys	Gly 350	Lys	Pro
Glu	Ile	Ala 355	Cys	Glu	Asn	Pro	His 360	Сув	Thr	Val	Ser	Thr 365	Phe	Glu	Ala
Ala	Tyr 370	Ser	Thr	His	Cys	Arg 375	Gln	Arg	Ser	Asn	Pro 380	Arg	Gly	Ala	Gly
Ile 385	Arg	Ser	Leu	Суѕ	Arg 390	Xaa									

104

<210> 152

<211> 99

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (99)

<223> Xaa equals stop translation

<400> 152

Met Thr Thr Arg Gln Pro Thr Ala Val Ser Trp Pro Cys Trp Leu Met

1 5 10 15

Ser Ser Ser Leu Ser Thr Ala Cys Leu Ala Trp Thr Leu Thr Gly Ser 20 25 30

Leu Ala Arg Glu Ala Thr Arg Arg Ala Arg Ser Leu Ser Pro Thr Trp
35 40 45

Asn Cys Ser Ala Arg Gln Val Pro Pro Ser Pro Pro His Ser Gly Leu
50 55 60

Gly Arg Arg Gly Trp Ala His Cys His Leu Thr Cys Leu Leu Val Thr
65 70 75 80

Gln Leu Phe Arg Val Gly Arg Ile His Pro Ile Leu Ser Leu Pro Leu 85 90 95

Val Thr Xaa

<210> 153

<211> 61

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (61)

<223> Xaa equals stop translation

<400> 153

Met Ser His Cys Ala Arg Pro Thr Phe Leu Thr Leu Leu Leu Ala Ser 1 5 10 15

Cys Phe Trp Ala Ala Ala Ile Pro Asn Arg Asn Val Ile Leu Ser Val 20 25 30

Ser Phe Arg Pro Leu His Met Gln Phe Thr Leu Ser Ile Leu Val Phe 35 40 45

Ile Leu Arg Ile Leu Ile Leu Leu Arg Ser Phe Leu Xaa 50 55 60

<210> 154

105

<211> 393 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (393) <223> Xaa equals stop translation <400> 154 Met Glu Trp Trp Ala Ser Ser Pro Leu Arg Leu Trp Leu Leu Phe 1.0 Leu Leu Pro Ser Ala Gln Gly Arg Gln Lys Glu Ser Gly Ser Lys Trp 25 Lys Val Phe Ile Asp Gln Ile Asn Arg Ser Leu Glu Asn Tyr Glu Pro Cys Ser Ser Gln Asn Cys Ser Cys Tyr His Gly Val Ile Glu Glu Asp Leu Thr Pro Phe Arg Gly Gly Ile Ser Arg Lys Met Met Ala Glu Val Val Arg Arg Lys Leu Gly Thr His Tyr Gln Ile Thr Lys Asn Arg Leu Tyr Arg Glu Asn Asp Cys Met Phe Pro Ser Arg Cys Ser Gly Val Glu 105 His Phe Ile Leu Glu Val Ile Gly Arg Leu Pro Asp Met Glu Met Val 120 125 Ile Asn Val Arg Asp Tyr Pro Gln Val Pro Lys Trp Met Glu Pro Ala 135 Ile Pro Val Phe Ser Phe Ser Lys Thr Ser Glu Tyr His Asp Ile Met 155 Tyr Pro Ala Trp Thr Phe Trp Glu Gly Gly Pro Ala Val Trp Pro Ile 165 Tyr Pro Thr Gly Leu Gly Arg Trp Asp Leu Phe Arg Glu Asp Leu Val 185 Arg Ser Ala Ala Gln Trp Pro Trp Lys Lys Lys Asn Ser Thr Ala Tyr 200 195 Phe Arg Gly Ser Arg Thr Ser Pro Glu Arg Asp Pro Leu Ile Leu Leu 220 215 Ser Arg Lys Asn Pro Lys Leu Val Asp Ala Glu Tyr Thr Lys Asn Gln Ala Trp Lys Ser Met Lys Asp Thr Leu Gly Lys Pro Ala Ala Lys Asp 250 245

Val His Leu Val Asp His Cys Lys Tyr Lys Tyr Leu Phe Asn Phe Arg 260 265 270

Gly Val Ala Ala Ser Phe Arg Phe Lys His Leu Phe Leu Cys Gly Ser 275 280 285

Leu Val Phe His Val Gly Asp Glu Trp Leu Glu Phe Phe Tyr Pro Gln 290 295 300

Leu Lys Pro Trp Val His Tyr Ile Pro Val Lys Thr Asp Leu Ser Asn 305 310 315 320

Val Gln Glu Leu Leu Gln Phe Val Lys Ala Asn Asp Asp Val Ala Gln 325 330 335

Glu Ile Ala Glu Arg Gly Ser Gln Phe Ile Arg Asn His Leu Gln Met 340 345 350

Asp Asp Ile Thr Cys Tyr Trp Glu Asn Leu Leu Ser Glu Tyr Ser Lys 355 360 365

Phe Leu Ser Tyr Asn Val Thr Arg Arg Lys Gly Tyr Asp Gln Ile Ile 370 375 380

Pro Lys Met Leu Lys Thr Glu Leu Xaa 385 390

<210> 155

<211> 75

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (75)

<223> Xaa equals stop translation

<400> 155

Met Leu Ile Leu Phe Leu Ser Val Cys Leu Phe Val Phe Leu Leu Thr 1 5 10 15

Val Arg Ala Leu Cys Cys Arg Ser Ala Gly Val Trp Leu Arg Ser Thr 20 25 30

Pro Asp Pro Val Cys Leu Gly Phe Ala Arg Gly Gly Cys Arg Ile Ala 35 40 45

Met Ile Ala Ala Cys Phe Ser Ser Gly Ser Phe Val Pro Glu Gly His 50 55 60

Pro Pro Asp Ala Ser Gln Ser Ser Pro Val Xaa 65 70 75

<210> 156

<211> 82

<212> PRT

```
<213> Homo sapiens
<220>
<221> SITE
<222> (82)
<223> Xaa equals stop translation
<400> 156
Met Trp Pro Leu Leu Ala Val Ser Pro Phe Gly Leu Val Trp Ala Ser
Ser Gln Ser Gly Ser Leu Leu Leu Arg Ala Ser Ile Pro Arg Gln His
Ser Arg Arg Ala Trp His Phe Tyr Ser Glu Val Trp Gln Ser His Ser
                             40
Val Ala Ser Val Leu Leu Tyr Leu Leu Val Arg Ala Ile Thr Lys Met
     50
Cys Ile Gly Ser Lys Lys Arg Asp Ile Thr Pro Thr Thr Arg Trp Lys
 65
Lys Xaa
<210> 157
<211> 54
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (49)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (54)
<223> Xaa equals stop translation
<400> 157
Met Ser His His Ala Gly Leu Gly Gly Gly Ile Leu Phe Ser Leu Lys
Ile Ser Phe Phe Ile Ala Leu Ala Val Val Gly Gly Ser Arg Gly Val
             20
Asn Asp Cys Gln Leu Gly Gly Cys Arg Val Gly Ser Cys Pro Arg Val
Xaa Val Arg Val Ala Xaa
     50
<210> 158
<211> 103
```

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (103)

<223> Xaa equals stop translation

<400> 158

Met Thr Val Arg Arg Leu Ser Leu Leu Cys Arg Asp Leu Trp Ala Leu
1 5 10 15

Trp Leu Leu Lys Ala Gly Ala Val Arg Gly Ala Arg Ala Gly Pro 20 25 30

Arg Leu Pro Gly Arg Cys Cys Gly Ala Thr Cys Gly Asp Ala Gly Arg
35 40 45

Gly Trp Thr Phe Trp Ala Gln Pro Cys Pro Gln Lys Leu Leu Gly Gln 50 60

Lys Pro Gly Ala Gly Gly Cys Arg Gly Trp Val Leu Gly Trp Val Pro 65 70 75 80

Pro Arg Pro Glu Glu Pro Cys Ser Leu Ala Gly Lys Val Cys Thr Gly 85 90 95

Leu Ala Arg Trp Met Val Xaa 100

<210> 159

<211> 575

<212> PRT

<213> Homo sapiens

<400> 159

Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala 1 5 10 15

Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys 20 25 30

Phe Arg Gln Glu Ser His Val Glu Phe Ala Thr Gly Glu Leu Leu Thr 35 40 45

Met Thr Gln Trp Pro Gly Val Trp Ile Pro Met Ala Ser Cys Ser Ser 50 55 60

Thr Trp Trp Ser Met Ala Leu Ser Pro Asp Ser Leu Ala Asp Ala Asp 65 70 75 80

Leu Gln Val Gln Asp Phe Glu Glu His Tyr Val Gln Thr Gly Pro Gly 85 90 95

Gln Leu Phe Val Gly Ser Thr Gln Arg Phe Phe Gln Gly Gly Leu Pro 100 105 110

Ser 	Phe	Leu 115	Arg	Cys	Asn	His	Ser 120	Ile	Gln	Tyr	Asn	Ala 125	Ala	Arg	Gly
Pro	Gln 130	Pro	Gln	Leu	Val	Gln 135	His	Leu	Arg	Ala	Ser 140	Ala	Ile	Ser	Ser
Ala 145	Phe	Asp	Pro	Glu	Ala 150	Glu	Ala	Leu	Arg	Phe 155	Gln	Leu	Ala	Thr	Ala 160
Leu	Gln	Ala	Glu	Glu 165	Asn	Glu	Val	Gly	Cys 170	Pro	Glu	Gly	Phe	Glu 175	Leu
Asp	Ser	Gln	Gly 180	Ala	Phe	Cys	Val	Asp 185	Val	Asp	Glu	Cys	Ala 190	Trp	Asp
Ala	His	Leu 195	Cys	Arg	Glu	Gly	Gln 200	Arg	Cys	Val	Asn	Leu 205	Leu	Gly	Ser
Tyr	Arg 210	Cys	Leu	Pro	Asp	Cys 215	Gly	Pro	Gly	Phe	Arg 220	Val	Ala	Asp	Gly
Ala 225	Gly	Cys	Glu	Asp	Val 230	Asp	Glu	Cys	Leu	Glu 235	Gly	Leu	Asp	Asp	Суs 240
His	Tyr	Asn	Gln	Leu 245	Cys	Glu	Asn	Thr	Pro 250	Gly	Gly	His	Arg	Cys 255	Ser
Суѕ	Pro	Arg	Gly 260	Tyr	Arg	Met	Gln	Gly 265	Pro	Ser	Leu	Pro	Cys 270	Leu	Asp
Val	Asn	Glu 275	Cys	Leu	Gln	Leu	Pro 280	Lys	Ala	Cys	Ala	Tyr 285	Gln	Cys	His
Asn	Leu 290	Gln	Gly	Ser	Tyr	Arg 295	Cys	Leu	Cys	Pro	Pro 300	Gly	Gln	Thr	Leu
Leu 305	Arg	Asp	Gly	Lys	Ala 310	Cys	Thr	Ser	Leu	Glu 315	Arg	Asn	Gly	Gln	Asn 320
Val	Thr	Thr	Val	Ser 325	His	Arg	Gly	Pro	Leu 330	Leu	Pro	Trp	Leu	Arg 335	Pro
Trp	Ala	Ser	11e 340	Pro	Gly	Thr	Ser	Tyr 345	His	Ala	Trp	Val	Ser 350	Leu	Arg
Pro	Gly	Pro 355	Met	Ala	Leu	Ser	Ser 360	Val	Gly	Arg	Ala	Trp 365	Cys	Pro	Pro
Gly	Phe 370	Ile	Arg	Gln	Asn	Gly 375	Val	Cys	Thr	Asp	Leu 380	Asp	Glu	Суѕ	Arg
Val 385	Arg	Asn	Leu	Cys	Gln 390	His	Ala	Суѕ	Arg	Asn 395	Thr	Glu	Gly	Ser	Tyr 400
Gln	Cys	Leu	Cys	Pro 405	Ala	Gly	Tyr	Arg	Leu 410	Leu	Pro	Ser	Gly	Lys 415	Asn
Cys	Gln	Asp	Ile	Asn	Glu	Cys	Glu	Glu	Glu	Ser	Ile	Glu	Cys	Gly	Pro

110

420 425 430 Gly Gln Met Cys Phe Asn Thr Arg Gly Ser Tyr Gln Cys Val Asp Thr 440 Pro Cys Pro Ala Thr Tyr Arg Gln Gly Pro Ser Pro Gly Thr Cys Phe 455 Arg Arg Cys Ser Gln Asp Cys Gly Thr Gly Gly Pro Ser Thr Leu Gln 475 Tyr Arg Leu Leu Pro Leu Pro Leu Gly Val Arg Ala His His Asp Val 485 490 Ala Arg Leu Thr Ala Phe Ser Glu Val Gly Val Pro Ala Asn Arg Thr 505 Glu Leu Ser Met Leu Glu Pro Asp Pro Arg Ser Pro Phe Ala Leu Arg Pro Leu Arg Ala Gly Leu Gly Ala Val Tyr Thr Arg Arg Ala Leu Thr 535 Arg Ala Gly Leu Tyr Arg Leu Thr Val Arg Ala Ala Ala Pro Arg His 545 550 555 Gln Ser Val Phe Val Leu Leu Ile Ala Val Ser Pro Tyr Pro Tyr <210> 160 <211> 643 <212> PRT <213> Homo sapiens <400> 160 Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu 25 Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu 65 70 Pro Ser Cys Thr Asp Gln Cys Phe Arg Gln Asn His Ile His Gly Thr Leu Leu Gln Val Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys His 100 105 110 Gly Thr Asn Ser Asp Phe Gln His Glu Leu Thr Asp Ile Thr Val Cys

- Thr Lys Ala Lys Leu Trp Leu Ala Lys Arg Gln Asn Pro Cys Leu Val 130 135 140
- Thr Arg Ala Val Tyr Ile Asp Ile Leu Phe Leu Leu Thr Cys Cys Leu 155 160
- Asn Arg Ser Ala Lys Asp Asn Gln Pro Val Leu Glu Ser Leu Gly Phe
 165 170 175
- Trp Glu Glu Val Arg Gly Ile Ile Ser Gly Ser Glu Leu Ile Thr Gly
 180 185 190
- Phe Pro Trp Ala Phe Lys Val Pro Gly Leu Pro Gln Tyr Leu Gln Ser 195 200 205
- Leu Thr Arg Leu Ala Ile Ala Ala Val Trp Ala Ala Ala Ala Lys Ser 210 215 220
- Gly Glu Arg Glu Thr Asn Val Pro Ile Ser Phe Ser Gln Leu Leu Glu 225 230 235 240
- Ser Ala Phe Pro Glu Val Arg Ser Leu Thr Leu Glu Ala Leu Leu Glu 245 250 255
- Lys Phe Leu Ala Ala Ala Ser Gly Leu Gly Glu Lys Gly Val Pro Pro 260 265 270
- Leu Leu Cys Asn Met Gly Glu Lys Phe Leu Leu Leu Ala Met Lys Glu 275 280 285
- Asn His Pro Glu Cys Phe Cys Lys Ile Leu Lys Ile Leu His Cys Met 290 295 300
- Asp Pro Gly Glu Trp Leu Pro Gln Thr Glu His Cys Val His Leu Thr 305 310 315 320
- Pro Lys Glu Phe Leu Ile Trp Thr Met Asp Ile Ala Ser Asn Glu Arg
 325 330 335
- Ser Glu Ile Gln Ser Val Ala Leu Arg Leu Ala Ser Lys Val Ile Ser 340 345 350
- His His Met Gln Thr Cys Val Glu Asn Arg Glu Leu Ile Ala Ala Glu 355 360 365
- Leu Lys Gln Trp Val Gln Leu Val Ile Leu Ser Cys Glu Asp His Leu 370 375 380
- Pro Thr Glu Ser Arg Leu Ala Val Val Glu Val Leu Thr Ser Thr Thr 385 390 395 400
- Pro Leu Phe Leu Thr Asn Pro His Pro Ile Leu Glu Leu Gln Asp Thr 405 410 415
- Leu Ala Leu Trp Lys Cys Val Leu Thr Leu Leu Gln Ser Glu Glu Gln 420 425 430

112

Ala Val Arg Asp Ala Ala Thr Glu Thr Val Thr Thr Ala Met Ser Gln 440 Glu Asn Thr Cys Gln Ser Thr Glu Phe Ala Phe Cys Gln Val Asp Ala 455 Ser Ile Ala Leu Ala Leu Ala Leu Ala Val Leu Cys Asp Leu Leu Gln 470 Gln Trp Asp Gln Leu Ala Pro Gly Leu Pro Ile Leu Leu Gly Trp Leu Leu Gly Glu Ser Asp Asp Leu Val Ala Cys Val Glu Ser Met His Gln 500 505 Val Glu Glu Asp Tyr Leu Phe Glu Lys Ala Glu Val Asn Phe Trp Ala 520 Glu Thr Leu Ile Phe Val Lys Tyr Leu Cys Lys His Leu Phe Cys Leu Leu Ser Lys Ser Gly Trp Arg Pro Pro Ser Pro Glu Met Leu Cys His Leu Gln Arg Met Val Ser Glu Gln Cys His Leu Leu Ser Gln Phe Phe 570 Arg Glu Leu Pro Pro Ala Ala Glu Phe Val Lys Thr Val Glu Phe Thr 585 580 Arg Leu Arg Ile Gln Glu Glu Arg Thr Leu Ala Cys Leu Arg Leu Leu 600 Ala Phe Leu Glu Gly Lys Glu Gly Glu Asp Thr Leu Val Leu Ser Val 610 Trp Asp Ser Tyr Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu 630 635 Ala Ala Cys <210> 161 <211> 191 <212> PRT <213> Homo sapiens <220> <221> SITE

<222> (191)

<223> Xaa equals stop translation

113

Leu Leu Gly Val Ala Ala Ser Leu Cys Val Arg Cys Ser Arg Pro Gly
20 25 30

Ala Lys Arg Ser Glu Lys Ile Tyr Gln Gln Arg Ser Leu Arg Glu Asp 35 40 45

Gln Gln Ser Phe Thr Gly Ser Arg Thr Tyr Ser Leu Val Gly Gln Ala
50 55 60

Trp Pro Gly Pro Leu Ala Asp Met Ala Pro Thr Arg Lys Asp Lys Leu 65 70 75 80

Leu Gln Phe Tyr Pro Ser Leu Glu Asp Pro Ala Ser Ser Arg Tyr Gln 85 90 95

Asn Phe Ser Lys Gly Ser Arg His Gly Ser Glu Glu Ala Tyr Ile Asp 100 105 110

Pro Ile Ala Met Glu Tyr Tyr Asn Trp Gly Arg Phe Ser Lys Pro Pro 115 120 125

Glu Asp Asp Asp Ala Asn Ser Tyr Glu Asn Val Leu Ile Cys Lys Gln 130 135 140

Lys Thr Thr Glu Thr Gly Ala Gln Gln Glu Gly Ile Gly Gly Leu Cys 145 150 155 160

Arg Gly Asp Leu Ser Leu Ser Leu Ala Leu Lys Thr Gly Pro Thr Ser 165 170 175

Gly Leu Cys Pro Ser Ala Ser Pro Glu Glu Asp Glu Gly Ile Xaa 180 185 190

<210> 162

<211> 64

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (64)

<223> Xaa equals stop translation

<400> 162

Met Lys His Val Leu Asn Leu Tyr Leu Leu Gly Val Val Leu Thr Leu

1 5 10 15

Leu Ser Ile Phe Val Arg Val Met Glu Ser Leu Glu Gly Leu Leu Glu
20 25 30

Ser Pro Ser Pro Gly Thr Ser Trp Thr Thr Arg Ser Gln Leu Ala Asn
35 40 45

Thr Glu Pro Thr Lys Gly Leu Pro Asp His Pro Ser Arg Ser Met Xaa 50 55 60

114

```
<210> 163
 <211> 118
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (118)
 <223> Xaa equals stop translation
 <400> 163
 Met Ile Phe Leu Thr Val Leu Pro Leu Ala Phe Leu Phe Leu His Ser
 Gly Phe Tyr His Tyr Ile Ser Phe Ser Cys Leu Phe Ser Leu Ser Leu
                                  25
              20
 Ala Leu Phe Phe Leu Asp Val Ala Thr Phe Arg Arg Pro Gly Gln
                              40
 Leu Phe Cys Glu Arg Ser Val Leu Phe Asp Met Phe His Phe Gly Phe
 Val Ser Leu Phe Leu His Glu Trp Ile Gln Ala Lys His Phe Trp Ala
 Gly Leu Phe Ile Val Leu Pro Ser Asp Val Phe Phe Ser Val His His
 Leu Glu Ala Pro Asp Gly Ser Phe Pro Asn Ile Ala Lys Leu Ser Leu
             100
                                 105
 Ile Ile Leu Leu Arg Xaa
         115
 <210> 164
 <211> 43
 <212> PRT
<213> Homo sapiens
 <220>
 <221> SITE
 <222> (43)
```

<400> 164

Met Leu Gln Phe Thr Leu Trp Val Phe Gly Ala Ile His Phe Pro 1 5 10 15

Lys Cys Leu Gly Ile Lys Glu Glu Leu Leu Lys Cys Cys Leu Gln Leu 20 25 30

Pro Pro Ser Ser Thr Tyr Glu Lys Val Val Xaa

<223> Xaa equals stop translation

115

```
<210> 165
<211> 48
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (48)
<223> Xaa equals stop translation
<400> 165
Met Leu Ser Arg Arg Leu His Cys Leu Val Leu Tyr Phe Leu Leu Leu
Leu Leu Ser Phe Ile His Thr Leu Ser Val Ser His Ile Cys Ser Ser
                                 25
Phe Ile Trp Leu Phe Pro Lys Asn Ile Glu Ser Glu Ala Thr Met Xaa
                            40
<210> 166
<211> 46
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (46)
<223> Xaa equals stop translation
<400> 166
Met Glu Lys Met Gly Gln Gly Leu Leu Ser Ser Thr Tyr Leu Thr Val
```

Lys Glu Ser Lys Tyr Leu Ile Lys Thr Leu Gly Ser Gly Xaa

Leu His Leu Ile Gln Leu Val Gly Cys Gly Leu Leu Thr Glu Glu Ile

15 did Ser Lys Tyr Leu lie Lys Thr Leu Gly Ser Gly Xaa 35 40 45

<210> 167 <211> 207 <212> PRT <213> Homo sapiens

<400> 167
Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn Cys Ile Phe Phe
1 5 10 15

Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser 20 25 30

116

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Pro Leu 35 40 45

Pro Ala Cys Leu Asn Pro Val Leu Tyr Val Phe Phe Asn Pro Lys Phe 50 55 60

Lys Glu Asp Trp Lys Leu Leu Lys Arg Arg Val Thr Lys Lys Ser Gly 65 70 75 80

Ser Val Ser Val Ser Ile Ser Ser Gln Gly Gly Cys Leu Glu Gln Asp 85 90 95

Phe Tyr Tyr Asp Cys Gly Met Tyr Ser His Leu Gln Gly Asn Leu Thr 100 105 110

Val Cys Asp Cys Cys Glu Ser Phe Leu Leu Thr Lys Pro Val Ser Cys 115 120 125

Lys His Leu Ile Lys Ser His Ser Cys Pro Ala Leu Ala Val Ala Ser 130 135 140

Cys Gln Arg Pro Glu Gly Tyr Trp Ser Asp Cys Gly Thr Gln Ser Ala 145 150 155 160

His Ser Asp Tyr Ala Asp Glu Glu Asp Ser Phe Val Ser Asp Ser Ser 165 170 175

Asp Gln Val Gln Ala Cys Gly Arg Ala Cys Phe Tyr Gln Ser Arg Gly 180 185 190

Phe Pro Leu Val Arg Tyr Ala Tyr Asn Leu Pro Arg Val Lys Asp 195 · 200 205

<210> 168

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals stop translation

<400> 168

Met Tyr Ile Phe Glu Leu Ser Leu Tyr Leu Glu Gly Thr Ser Phe Val 1 5 10 15

Val Val Leu Leu Phe Leu Leu Ile Ser Val Ser Leu Asp Ser Pro Pro 20 25 30

Thr Thr Lys Gly Trp Asp Ser Val Leu His Ile Trp Val Pro Leu Ile 35 40 45

Val Gln Xaa

```
<210> 169
<211> 43
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (43)
<223> Xaa equals stop translation
<400> 169
Met Ala His Pro Gly Leu Pro Lys Thr Val Pro Val Tyr Ala Val Val
                                    10
Leu Ala Leu Leu Ile Met Thr Leu Pro Leu Thr Leu Thr Ile Asn Leu
             20
Asp Asp Asn Leu Tyr Gly Asn Ser Ala Lys Xaa
<210> 170
<211> 56
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation
<400> 170
Met Arg Pro Trp Trp Ser Leu Leu Leu Glu Ala Cys Ala Thr Cys Ala
Gln Thr Gly Pro Thr Arg Ser Thr Ser Cys Thr Gln Glu Val Ser His
             20
Ser Ser Ser Thr Ala Tyr Pro Ala Pro Met Arg Arg Cys Cys Leu
                             40
Pro Ser Pro Arg Ser Cys Thr Xaa
<210> 171
<211> 109
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (109)
<223> Xaa equals stop translation
<400> 171
Met Ala Leu Ala Gly Ser Val Phe Val Leu Gly Gly Val Leu Val Leu
```

118 10 Cys Val Glu Arg Asn Gly Glu Gly Glu Met Gly Trp Pro Gln His Leu 25 Pro Lys Ser Gln Pro Leu Ser Pro Pro Val Ala Val Arg Arg Cys Ser 40 Phe Glu Arg Ser Trp Ile Asp Leu Leu Val Glu Thr Ser Ser Met Val Thr Cys Arg Gln Gln Val Gly Thr Pro Asn Gly Met Glu Gly Arg 65 Gly Gly Pro Lys Thr Thr Phe Pro Ile Arg Leu Gln Leu Ser Gly 85 Ala Cys Ala Val Arg Pro Glu Ile Gln Trp Glu Val Xaa 1.05 <210> 172 <211> 51 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (17) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (51) <223> Xaa equals stop translation <400> 172 Met Phe Leu Phe Phe Tyr Leu Ser Leu Ala Val Tyr Ala Gln Arg Gln Xaa Ser Gly Ser Cys Arg Gln Thr Asp His Arg Trp Lys Ser Arg Gly 20 Ala Arg Arg Cys Phe Leu Glu Pro Arg Asp Pro Gly Ser Val Pro Gly 40 His Pro Xaa 50 <210> 173 <211> 566 <212> PRT <213> Homo sapiens <400> 173

Met Ala Pro Leu Ala Leu His Leu Leu Val Leu Val Pro Ile Leu Leu

Ser	Leu	Val	Ala 20	Ser	Gln	Asp	Trp	Lys 25	Ala	Glu	Arg	Ser	30	Asp	Pro
Phe	Glu	Lys 35	Cys	Met	Gln	Asp	Pro 40	Asp	Туr	Glu	Gln	Leu 45	Leu	Lys	Val
Val	Thr 50	Trp	Gly	Leu	Asn	Arg 55	Thr	Leu	Lys	Pro	Gln 60	Arg	Val	Ile	Val
Val 65	Gly	Ala	Gly	Val	Ala 70	Gly	Leu	Val	Ala	Ala 75	Lys	Val	Leu	Ser	Asp 80
Ala	Gly	His	Lys	Val 85	Thr	Ile	Leu	Glu	Ala ·90	Asp	Asn	Arg	Ile	Gly 95	Gly
Arg	Ile	Phe	Thr 100	Tyr	Arg	Asp	Gln	Asn 105	Thr	Gly	Trp	Ile	Gly 110	Glu	Leu
Gly	Ala	Met 115	Arg	Met	Pro	Ser	Ser 120	His	Arg	Ile	Leu	His 125	Lys	Leu	Суѕ
Gln	Gly 130	Leu	Gly	Leu	Asn	Leu 135	Thr	Lys	Phe	Thr	Gln 140	Tyr	Asp	Lys	Asn
Thr 145	Trp	Thr	Glu	Val	His 150	Glu	Val	Lys	Leu	Arg 155	Asn	Tyr	Val	Val	Glu 160
Lys	Val	Pro	Glu	Lys 165	Leu	Gly	Tyr	Ala	Leu 170	Arg	Pro	Gln	Glu	Lys 175	Gly
His	Ser	Pro	Glu 180	Asp	Ile	Tyr	Gln	Met 185	Ala	Leu	Asn	Gln	Ala 190	Leu	Lys
Asp	Leu	Lys 195	Ala	Leu	Gly	Cys	Arg 200	Lys	Ala	Met	Lys	Lys 205	Phe	Glu	Arg
His	Thr 210	Leu	Leu	Glu	Tyr	Leu 215	Leu	Gly	Glu	Gly	Asn 220	Leu	Ser	Arg	Pro
225					230					235			Phe		240
				245					250				Ser	255	
Leu	Gln	Tyr	Ser 260	Arg	Ile	Val	Gly	Gly 265	Trp	Asp	Leu	Leu	Pro 270	Arg	Ala
Leu	Leu	Ser 275	Ser	Leu	Ser	Gly	Leu 280	Val	Leu	Leu	Asn	Ala 285	Pro	Val	Val
Ala	Met 290	Thr	Gln	Gly	Pro	His 295	Asp	Val	His	Val	Gln 300		Glu	Thr	Ser
Pro 305	Pro	Ala	Arg	Asn	Leu 310	Lys	Val	Leu	Lys	Ala 315	Asp	Val	Val	Leu	Leu 320

120

Thr Ala Ser Gly Pro Ala Val Lys Arg Ile Thr Phe Ser Pro Pro Leu 325 330 335

Pro Arg His Met Gln Glu Ala Leu Arg Arg Leu His Tyr Val Pro Ala 340 345 350

Thr Lys Val Phe Leu Ser Phe Arg Arg Pro Phe Trp Arg Glu Glu His 355 360 365

Ile Glu Gly Gly His Ser Asn Thr Asp Arg Pro Ser Arg Met Ile Phe 370 380

Tyr Pro Pro Pro Arg Glu Gly Ala Leu Leu Leu Ala Ser Tyr Thr Trp 385 390 395 400

Ser Asp Ala Ala Ala Ala Phe Ala Gly Leu Ser Arg Glu Glu Ala Leu 405 410 415

Arg Leu Ala Leu Asp Asp Val Ala Ala Leu His Gly Pro Val Val Arg
420 425 430

Gln Leu Trp Asp Gly Thr Gly Val Val Lys Arg Trp Ala Glu Asp Gln 435 440 445

His Ser Gln Gly Gly Phe Val Val Gln Pro Pro Ala Leu Trp Gln Thr 450 455 460

Glu Lys Asp Asp Trp Thr Val Pro Tyr Gly Arg Ile Tyr Phe Ala Gly 465 470 475 480

Glu His Thr Ala Tyr Pro His Gly Trp Val Glu Thr Ala Val Lys Leu 485 490 495

Leu Arg Ala Ala Ile Lys Ile Asn Ser Arg Lys Gly Pro Ala Ser Asp 500 505 510

Thr Ala Ser Pro Glu Gly His Ala Ser Asp Met Glu Gly Gln Gly His 515 520 525

Val His Gly Val Ala Ser Ser Pro Ser His Asp Leu Ala Lys Glu Glu 530 540

Gly Ser His Pro Pro Val Gln Gly Gln Leu Ser Leu Gln Asn Thr Thr 545 550 555 560

His Thr Arg Thr Ser His 565

<210> 174

<211> 224

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (76)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 174 Met Ala Arg Ala Arg Gly Ser Pro Cys Pro Pro Leu Pro Pro Gly Arg 10 Met Ser Trp Pro His Gly Ala Leu Leu Phe Leu Trp Leu Phe Ser Pro Pro Leu Gly Ala Gly Gly Gly Val Ala Val Thr Ser Ala Ala Gly Gly Gly Ser Pro Pro Ala Thr Ser Cys Pro Val Ala Cys Ser Cys Ser Asn Gln Ala Ser Arg Val Ile Cys Thr Arg Arg Xaa Leu Ala Glu Val 75 Pro Ala Ser Ile Pro Val Asn Thr Arg Tyr Leu Asn Leu Gln Glu Asn Gly Ile Gln Val Ile Arg Thr Asp Thr Phe Lys His Leu Arg His Leu 105 Glu Ile Leu Gln Leu Ser Lys Asn Leu Val Arg Lys Ile Glu Val Gly 120 Ala Phe Asn Gly Leu Pro Ser Leu Asn Thr Leu Glu Leu Phe Asp Asn 135 Arg Leu Thr Thr Val Pro Thr Gln Ala Phe Glu Tyr Leu Ser Lys Leu 150 155 Arg Glu Leu Trp Leu Arg Asn Asn Pro Ile Glu Ser Ile Pro Ser Tyr Ala Phe Asn Arg Val Pro Ser Leu Arg Arg Leu Asp Leu Gly Glu Leu 185 Lys Arg Leu Glu Tyr Ile Ser Glu Ala Ala Phe Glu Gly Leu Val Asn 195 Leu Arg Tyr Leu Asn Leu Gly Met Cys Asn Leu Lys Asp Ile Pro Asn

<210> 175 <211> 123

210

<212> PRT

<213> Homo sapiens

<400> 175

Met His Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys Pro Ser 1 5 10 15

215

220

Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Val His Leu Asn Ile

122

20 25 30

Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr 35 40 45

Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe Gln Tyr
50 55 60

Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly Leu Thr
65 70 75 80

Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe Asp Ser 85 90 95

Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg Ala Ala 100 105 110

Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu 115 120

<210> 176

<211> 60

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (60)

<223> Xaa equals stop translation

<400> 176

Met Gly Leu Ser Val Leu Leu Pro Leu Cys Leu Leu Gly Pro Gly Arg
1 5 10 15

Phe Thr Ser Gly Glm Lys Pro Leu Asp Thr Pro Gly Leu Gly Ala Ala 20 25 30

Val Leu Ser Val Arg Lys Ala Gly Leu Lys Met Arg Ser His Leu Thr 35 40 45

Pro Ser Val Cys Thr Val Pro Ser Pro Gly Ser Xaa 50 55 60

<210> 177

<211> 105

<212> PRT

<213> Homo sapiens

<400> 177

Met Asp Thr Val Phe Leu Ile Gln Tyr Leu Phe Leu Thr Phe Pro Arg

1 5 10 15

Ile Val Phe Met Leu Gly Phe Val Val Val Leu Ser Phe Leu Leu Gly
20 25 30

Gly Tyr Leu Leu Phe Val Leu Tyr Leu Ala Ala Thr Asn Gln Thr Thr

123

35 40 45

Asn Glu Trp Tyr Arg Gly Asp Trp Ala Trp Cys Gln Arg Cys Pro Leu 50 55 60

Val Ala Trp Pro Pro Ser Ala Glu Pro Gln Val His Arg Asn Ile His 65 70 75 80

Ser His Gly Leu Arg Ser Asn Leu Gln Glu Ile Phe Leu Pro Ala Phe 85 90 95

Pro Cys His Glu Arg Lys Lys Gln Glu 100 105

<210> 178

<211> 88

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (88)

<223> Xaa equals stop translation

<400> 178

Met Ala Asp Pro His Val Ser Phe Leu Ser Phe Arg Gln Leu Phe Ser 1 5 10 15

Trp Ala Ala Val Ile Leu Leu Arg Gly Ile Leu Gly Thr Val Ala Pro 20 25 30

Pro Pro Cys Pro Cys Val Leu Asp Leu Ala Val Tyr Pro Leu His Leu 35 40 45

Pro Val Glu Ala Pro Cys Leu Glu Val Val Phe Lys Gln Lys Asn Gly 50 55 60

Lys Asp Asn Cys Leu Val Phe Tyr Pro Asp Pro Ile Pro Leu Arg Gly 65 70 75 80

Ser Leu Leu Gly Pro Phe Ile Xaa 85

<210> 179

<211> 88

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (55)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (66)

124

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (88)

<223> Xaa equals stop translation

<400> 179

Met Ala Asp Pro His Val Ser Phe Leu Ser Phe Arg Gln Leu Phe Ser 1 5 10 15

Trp Ala Ala Val Ile Leu Leu Arg Gly Ile Leu Gly Thr Val Ala Pro 20 25 30

Pro Pro Cys Pro Cys Val Leu Asp Leu Ala Val Tyr Pro Leu His Leu
35 40 45

Pro Val Glu Ala Pro Cys Xaa Glu Val Val Phe Lys Gln Lys Asn Gly 50 55 60

Lys Xaa Asn Cys Leu Val Phe Tyr Pro Asp Pro Ile Pro Leu Arg Gly 65 70 75 80

Ser Leu Leu Gly Pro Phe Ile Xaa

<210> 180

<211> 49

<212> PRT

<213> Homo sapiens

<400> 180

Met Asn Leu Cly Met Ile Phe Ser Met Cys Gly Leu Met Leu Lys
1 10 15

Leu Lys Trp Cys Ala Trp Val Ala Val Tyr Cys Ser Phe Ile Ser Phe 20 25 30

Ala Asn Ser Arg Ser Ser Glu Asp Thr Lys Gln Met Met Ser Ser Phe 35 40 45

Met

<210> 181

<211> 23

<212> PRT

<213> Homo sapiens

<400> 181

Leu Gly Ser Leu Ser Thr Ala Pro Ser Ser Ala Leu Pro Thr Leu Gly
1 5 10 15

Ala Arg Arg Thr Arg Ser Lys

```
<210> 182
<211> 104
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (104)
<223> Xaa equals stop translation
```

<400> 182

Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile 20

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe . 40

Ser Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp 60 50 55

Val Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn

Tyr Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg

Thr Arg Val Leu Phe Ile Tyr Xaa 100

<210> 183 <211> 198

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (29)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 183

Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu

His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln His Xaa Phe Asp Tyr 20

Lys Asp Glu Ser Gly Phe Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr 40

Thr Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr

Ile Pro Leu Val Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp

126

65 70 75 80

Phe Asp Asp Ala Asp Gln Leu Arg Ile Gly Asn Asp Gly Ile Phe Met 85 90 95

Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe Leu 105 110

Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile Ser 115 120 125 .

Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe Ser 130 135 140

Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val 145 150 155 160

Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr 165 170 175

Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr 180 185 190

Arg Val Leu Phe Ile Tyr 195

<210> 184 <211> 70 <212> PRT

<213> Homo sapiens

<400> 184

Met Leu Leu His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln Phe Ser 1 5 10 15

Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val 20 25 30

Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr 35 40 45

Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr 50 55 60

Arg Val Leu Phe Ile Tyr 65 70

<210> 185 <211> 82 <212> PRT <213> Homo sapiens

<400> 185

Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe 1 5 10 15

127

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile 20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe . 35 40 45

Ser Thr Tyr Phe Pro Ala Phe Met Asn Ser Leu Ser Arg Ser Lys Arg 50 55 60

Thr Pro Ala Gly Ser Glu Ser Arg Cys Arg Thr Gln Arg Asn Asn His 65 70 75 80

Leu Leu

<210> 186

<211> 45

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 186

Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu 1 5 10 15

His Leu Thr Ala Ala Phe Leu Gln Arg Ala His Xaa Ile Leu Thr Thr 20 25 30

Arg Met Ser Leu Gly Phe Gln Ser Pro His Leu Thr Met
35 40 45

<210> 187

<211> 34

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (34)

<223> Xaa equals stop translation

<400> 187

Met Thr Val Met Asp Pro Lys Gln Met Asn Val Ala Ala Ala Val Trp

1 5 10 15

Ala Val Val Ser Tyr Val Val Ala Asp Met Glu Glu Met Leu Pro Arg 20 25 30

Ser Xaa

128

<210> 188

<211> 232

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (232)

<223> Xaa equals stop translation

<400> 188

Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser 1 10 15

Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr
20 25 30

Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro 35 40 . 45

Ile Lys Lys Asn Ser Gly His Ile Tyr Asn Lys Asn Ile Ser Gln Lys
50 55 60

Asp Cys Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro 65 70 75 80

Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg 85 90 95

Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu 100 105 110

Gly Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro Ile 115 120 125

Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp Asp 130 135 140

Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu Ala 145 150 155 160

Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Gly Thr 165 170 175

Ala Ala Leu Glu Ala Ser Ser Pro Arg Ala Ala Lys Ser Leu Ser Leu 180 185 190

Thr Gly Met Leu Ser Ser Ala Asn Trp Gly Ile Glu Phe Lys Val Thr
195 200 205

Arg Lys Lys Gln Ala Asp Asn Trp Lys Gly Thr Asp Trp Val Leu Leu 210 215 220

Gly Phe Ile Leu Ile Pro Cys Xaa 225 230

129

<211> 457

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (457)

<223> Xaa equals stop translation

<400> 189

Met Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro
1 5 10 15

Leu Leu Ala Gly Leu Ala Leu Leu Gly Val Gly Pro Val Pro Ala Arg 20 25 30

Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly
.35 40 45

Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu 50 55 60

Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln 65 70 75 80

Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His
85 90 95

Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser
100 105 110

Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser 115 120 125

Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro 130 135 140

Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile 145 150 155 160

Met Asp Phe Asn Gly Asp Leu Ile Pro Asp Ile Phe Gly Ile Thr Asn 165 170 175

Glu Ser Asn Gln Pro Gln Ile Leu Leu Gly Gly Asn Leu Ser Trp His 180 185 190

Pro Ala Leu Thr Thr Ser Lys Met Arg Ile Pro His Ser His Ala 195 200 205

Phe Ile Asp Leu Thr Glu Asp Phe Thr Ala Asp Leu Phe Leu Thr Thr 210 215 220

Leu Asn Ala Thr Thr Ser Thr Phe Gln Phe Glu Ile Trp Glu Asn Leu 225 230 235 240

Asp Gly Asn Phe Ser Val Ser Thr Ile Leu Glu Lys Pro Gln Asn Met 245 250 255

130

Met Val Val Gly Gln Ser Ala Phe Ala Asp Phe Asp Gly Asp Gly His 260 265 270

Met Asp His Leu Leu Pro Gly Cys Glu Asp Lys Asn Cys Gln Lys Ser 275 280 285

Thr Ile Tyr Leu Val Arg Ser Gly Met Lys Gln Trp Val Pro Val Leu 290 295 300

Gln Asp Phe Ser Asn Lys Gly Thr Leu Trp Gly Phe Val Pro Phe Val 305 310 315 320

Asp Glu Gln Gln Pro Thr Glu Ile Pro Ile Pro Ile Thr Leu His Ile 325 330 335

Gly Asp Tyr Asn Met Asp Gly Tyr Pro Asp Ala Leu Val Ile Leu Lys
340 345 350

Asn Thr Ser Gly Ser Asn Gln Gln Ala Phe Leu Leu Glu Asn Val Pro 355 360 365

Cys Asn Asn Ala Ser Cys Glu Glu Ala Arg Arg Met Phe Lys Val Tyr 370 375 380

Trp Glu Leu Thr Asp Leu Asn Gln Ile Lys Asp Ala Met Val Ala Thr 385 390 395 400

Phe Phe Asp Ile Tyr Glu Asp Gly Ile Leu Asp Ile Val Val Leu Ser 405 410 415

Lys Gly Tyr Thr Lys Asn Asp Phe Ala Ile His Thr Leu Lys Asn Asn 420 425 430

Phe Glu Ala Asp Ala Tyr Phe Val Lys Val Ile Val Leu Ser Gly Leu 435 440 445

Cys Ser Asn Asp Cys Pro Arg Arg Xaa 450 455

<210> 190

<211> 185

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (185)

<223> Xaa equals stop translation

<400> 190

Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser 1 5 10 15

Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile 20 25 30

Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg

131

Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys
50 55 60

40

Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser

70

His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys
85 90 95

Pro Trp Val Gly Asn Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr
100 105 110

Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe
115 120 125

Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu 130 135 140

Thr Leu Lys Gly Asn Ser Trp Asn Cys Ser Arg Ser Pro His Leu Leu 145 150 155 160

Leu Tyr Thr Leu Val Arg Arg Gly Thr Asp Trp Ile Ser Tyr Phe Pro 165 170 175

Arg Gly Ser Gln Pro Asp Asn Gln Xaa 180 185

<210> 191

<211> 147

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (147)

<223> Xaa equals stop translation

<400> 191

Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala 1 5 10 15

Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys 20 25 30

Phe Arg Gln Ser His Thr Trp Ser Leu Leu Gln Gly Ala Ala His Asp
45

Asp Pro Val Ala Arg Gly Leu Asp Pro Asp Gly Leu Leu Leu Asp 50 55 60

Val Val Val Asn Gly Val Val Pro Gly Arg Ala Trp Leu Thr Gln Ile
65 70 75 80

Phe Lys Cys Arg Thr Leu Lys Lys His Tyr Val Gln Thr Arg Ala Trp
85 90 95

132

Pro Ala Val Arg Gly Leu His Thr Ala Leu Leu Pro Gly Arg Pro Pro 100 105 110

Leu Val Pro Thr Leu Gln Pro Gln His Pro Val Gln Arg Gly Pro Gly
115 120 125

Pro Pro Ala Pro Ala Gly Ala Ala Pro Ala Gly Leu Ser Tyr Gln Leu 130 135 140

Gly Leu Xaa 145

<210> 192

<211> 125

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (125)

<223> Xaa equals stop translation

<400> 192

Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val 1 5 10 15

Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu 20 25 30

Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro 35 40 45

Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val 50 60

Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu 65 70 75 80

Pro Ser Cys Thr Asp Gln Cys Phe Arg Ala Lys Pro His Ser Trp Gly 85 90 95

His Phe Ser Arg Phe Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys
100 105 110

Thr Arg Asn Glu Phe Arg Leu Pro Ala Arg Ala Asp Xaa 115 120 125

<210> 193

<211> 52

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (52)

133

<223> Xaa equals stop translation

<400> 193

Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn Cys Ile Phe Phe 1 5 10 15

Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser 20 25 30

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Pro Cys
35 40 45

Leu Leu Ala Xaa 50

<210> 194

<211> 320

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (68)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (115)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (213)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (320)

<223> Xaa equals stop translation

<400> 194

Met Ala Pro Leu Ala Leu His Leu Leu Val Leu Val Pro Ile Leu Leu 1 5 10 15

Ser Leu Val Ala Ser Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro 20 25 30

Phe Glu Lys Cys Met Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val 35 40 45

Thr Ile Leu Glu Ala Asp Asn Arg Ile Gly Gly Arg Ile Phe Thr Tyr
50 60

Arg Asp Gln Xaa Thr Gly Trp Ile Gly Glu Leu Gly Ala Met Arg Met 65 70 75 80

Pro Ser Ser His Arg Ile Leu His Lys Leu Cys Gln Gly Leu Gly Leu

85

90

95

Asn Leu Thr Lys Phe Thr Gln Tyr Asp Lys Asn Thr Trp Thr Glu Val 100 105 110

His Glu Xaa Lys Leu Arg Asn Tyr Val Val Glu Lys Val Pro Glu Lys
115 120 125

Leu Gly Tyr Ala Leu Arg Pro Gln Glu Lys Gly His Ser Pro Glu Asp 130 135 140

Ile Tyr Gln Met Ala Leu Asn Gln Ala Leu Lys Asp Leu Lys Ala Leu 145 150 155 160

Gly Cys Arg Lys Ala Met Lys Lys Phe Glu Arg His Thr Leu Leu Glu 165 170 175

Tyr Leu Leu Gly Glu Gly Asn Leu Ser Arg Pro Ala Val Gln Leu Leu 180 185 190

Gly Asp Val Met Ser Glu Asp Gly Phe Phe Tyr Leu Ser Phe Ala Glu 195 200 205

Ala Leu Arg Ala Xaa Ser Cys Leu Ser Asp Arg Leu Gln Tyr Ser Arg 210 215 220

Ile Val Gly Gly Trp Asp Leu Leu Pro Arg Ala Leu Leu Ser Ser Leu 225 230 235 240

Ser Gly Leu Val Leu Leu Asn Ala Pro Val Val Ala Met Thr Gln Gly 245 250 255

Pro His Asp Val His Val Gln Ile Glu Thr Ser Pro Pro Ala Arg Asn 260 265 270

Leu Lys Val Leu Lys Ala Asp Val Val Leu Leu Thr Ala Ser Gly Pro 275 . 280 . 285

Ala Val Lys Arg Ile Thr Phe Ser Pro Arg Cys Pro Ala Thr Cys Arg 290 295 300

Arg Arg Cys Gly Gly Cys Thr Thr Cys Arg Pro Pro Arg Cys Ser Xaa 305 310 315 320

<210> 195

<211> 130

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (38)

<223> Xaa equals any of the naturally occurring L-amino acids

```
135
<220>
<221> SITE
<222> (53)
<223> Xaa equals any of the naturally occurring L-amino acids
<400> 195
Pro Phe Cys Ser Gly Phe Phe Pro Ser Leu Trp Ile Tyr Leu Pro Phe
Ile Phe Asn Val Ser Asp Leu Trp Met Gly Ser Leu Ser Gly Cys Ala
                                 25
Leu Pro Phe Cys Leu Xaa Val Phe Phe Leu Thr Val Ser Pro Ser Ala
                             40
Val Gly Leu Leu Xaa Phe Ala Gly Gly Pro Leu Gln Thr Leu Phe Ala
Trp Val Ser Pro Val Glu Ala Ala Glu Gln Gln Arg Leu Pro Val
                                         75
Leu Ser Ser Gly Ser Phe Val Ser Glu Gly Thr Cys Gln Met Pro Ala
                                     90
Arg Ala Leu Leu Tyr Glu Val Ser Val Gly Pro Tyr Trp Glu Ile Pro
                                105
Pro Ser Gln Asp Thr Arg Arg Ser Gly Thr Tyr Leu Arg Arg Gln Ser
                            120
Asp Pro
    130
<210> 196
<211> 108
<212> PRT
<213> Homo sapiens
<400> 196
His Glu Gly Ser Cys Arg Ala Pro Gly Phe Ser Ala His Lys Gly Arg
Gly Cys Pro Ser Pro Arg Met Thr Leu Pro Ser Arg Ala Leu Ala Ser
                                 25
Leu Gly Val Gly Val Trp Gly Met Leu Arg Leu Asn Gln Val Thr Val
Ser Cys Gly Gly Ser Arg Trp Ser Ser Arg Val Ala Leu Gly Ala Phe
Ser Trp Val Cys Gly Val Ala Leu Val Leu Gln Pro Ser Gly Gly Gly
                     70
```

Leu Gly Leu Thr Ser Pro Ser Glu Gly Cys Trp Glu Gly Glu Leu Ala

85

90

136

Leu Ala Val Leu Arg Ala Pro Gly Gly Ser Pro Ser 100 105

<210> 197

<211> 104

<212> PRT

<213> Homo sapiens

<400> 197

Ile Pro Leu Thr Leu Pro Gly Ile Phe Leu Leu Ile Arg Leu Phe Trp
1 5 10 15

Arg Leu Gly Gln Ser Ile Cys Gly Pro Gly Lys Leu Val Leu Trp Pro
20 25 30

Gln Phe Cys Cys Gly Cys Ala Val Ile Ser Gly His Cys Val Pro Arg 35 40 45

Gly Met Pro Ser Ser Trp Leu Pro Gly Cys Phe Val Leu Leu Cys Leu
50 60

Val Ala Val Gly Cys Gln Leu Arg Glu Trp Gly Val Gly Gly Val Ser
65 70 75 80

Ala Val Gly Leu Leu Ala Leu Pro His Leu Gln Val Leu Gly Met Arg 85 90 95

Gly Arg Gly Leu Ile Ser Gly Gly 100

<210> 198

<211> 237

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (142)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 198

Gly Pro Ala Gly Lys Glu Ala Trp Ile Trp Ser Trp Leu Leu Pro Ser

1 5 10 15

Pro Gly Pro Ala Pro Leu Pro Ser Ala Ser Trp Gly Leu Cys Gly Asp 20 25 30

Ala Pro Arg Ala Ala Ala Arg Gly Pro Val Glu Pro Gly Ala Ala Arg
35 40 45

Met Ala Leu Leu Ser Arg Pro Ala Leu Thr Leu Leu Leu Leu Met 50 55 60

Ala Ala Val Val Arg Cys Gln Glu Gln Ala Gln Thr Thr Asp Trp Arg 65 70 75 80

137

Ala Thr Leu Lys Thr Ile Arg Asn Gly Val His Lys Ile Asp Thr Tyr 85 90 95

Leu Asn Ala Ala Leu Asp Leu Leu Gly Gly Glu Asp Gly Leu Cys Gln 100 105 110

Tyr Lys Cys Ser Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys 115 120 125

Pro Ser Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Xaa His Leu 130 135 140

Asn Ile Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg 145 150 155 160

Cys Tyr Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe 165 170 175

Gln Tyr Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly 180 185 190

Leu Thr Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe 195 200 205

Asp Ser Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg 210 215 220

Ala Ala Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu 225 230 235

<210> 199

<211> 8

<212> PRT

<213> Homo sapiens

<400> 199

Cys Cys Asn Gln His Asp Arg Cys

<210> 200 .

<211> 15

<212> PRT

<213> Homo sapiens

<400> 200

Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr Glu Thr
1 5 10 15

<210> 201

<211> 16

<212> PRT

<213> Homo sapiens

<400> 201

Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr Glu Thr Cys Gly

1

5

10

15

<210> 202 <211> 260 <212> PRT <213> Homo sapiens <400> 202 Gly Thr Ser Ser Ala Arg Pro Arg Gly Ala Leu Pro Gly Gly Ser Ala Pro Ser Ala Pro His Gly Gln Leu Pro Gly Arg Ala Gln Pro Ala Pro 20 Val Ser Gly Pro Pro Pro Thr Ser Gly Leu Cys His Phe Asp Pro Ala Ala Pro Trp Pro Leu Trp Pro Gly Pro Trp Gln Leu Pro Pro His Pro Gln Asp Trp Pro Ala His Pro Asp Ile Pro Gln Asp Trp Val Ser Phe Leu Arg Ser Phe Gly Gln Leu Thr Leu Cys Pro Arg Asn Gly Thr Val 90 Thr Gly Lys Trp Arg Gly Ser His Val Val Gly Leu Leu Thr Thr Leu 100 Asn Phe Gly Asp Gly Pro Asp Arg Asn Lys Thr Arg Thr Phe Gln Ala 120 Thr Val Leu Gly Ser Gln Met Gly Leu Lys Gly Ser Ser Ala Gly Gln 130 135 Leu Val Leu Ile Thr Ala Arg Val Thr Thr Glu Arg Thr Ala Gly Thr 150 155 Cys Leu Tyr Phe Ser Ala Val Pro Gly Ile Leu Pro Ser Ser Gln Pro 170 Pro Ile Ser Cys Ser Glu Glu Gly Ala Gly Asn Ala Thr Leu Ser Pro Arg Met Gly Glu Cys Val Ser Val Trp Ser His Glu Gly Leu Val 195

Leu Thr Lys Leu Leu Thr Ser Glu Glu Leu Ala Leu Cys Gly Ser Arg

Leu Leu Val Leu Gly Ser Phe Leu Leu Phe Cys Gly Leu Leu Cys

Cys Val Thr Ala Met Cys Phe His Pro Arg Arg Glu Ser His Trp Ser

235

230

139

245 250 255

Arg Thr Arg Leu 260

<210> 203

<211> 80

<212> PRT

<213> Homo sapiens

<400> 203

Ala Arg Ala Pro Pro Gly Pro Glu Gly Leu Ser Pro Glu Ala Gln Pro 1 5 10 15

Pro Leu Leu Pro Met Gly Asn Cys Gln Ala Gly His Asn Leu His Leu 20 25 30

Cys Leu Ala His His Pro Pro Leu Val Cys Ala Thr Leu Ile Leu Leu 35 40 45

Leu Leu Gly Leu Ser Gly Leu Gly Leu Gly Ser Phe Leu Leu Thr His 50 55 60

Arg Thr Gly Leu Arg Thr Leu Thr Ser Pro Arg Thr Gly Ser Leu Phe 65 70 75 80

<210> 204

<211> 224

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (6)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (9)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (22)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (143)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

140

<222> (186)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 204

Arg Phe Leu Ser Val Xaa Pro Gln Xaa Glu Val Pro Phe Leu Leu His 1 5 10 15

Pro Cys Val Cys Phe Xaa Gly Gly His Pro Ser Leu Leu Pro Asp Pro 20 25 30

Cys Arg Ala Val Gly Gly Gly Trp Glu Ala Pro Arg Cys Cys Leu His

Glu Ala Leu Cys Gln Ser Leu Gly Cys Lys Ala Glu Glu Ile Val Ser · 50 55 60

Val Ser Glu Ser Ser Ser Ala Gln Arg Cys Trp Tyr Leu Leu Arg Gly 65 70 75 80

Arg Lys Ala Gly Gly Arg Gly Pro Ala Ser Pro Val Leu Phe Ala Leu 85 90 95

Met Arg Leu Glu Ser Leu Cys His Leu Cys Leu Ala Cys Leu Phe Phe 100 105 110

Arg Leu Pro Ala Thr Arg Thr Val Tyr Cys Met Asn Glu Ala Glu Ile 115 120 125

Val Asp Val Ala Leu Gly Ile Leu Ile Glu Ser Arg Lys Gln Xaa Lys 130 135 140

Ala Cys Glu Gln Pro Ala Leu Ala Gly Ala Asp Asn Pro Glu His Ser 145 150 155 160

Pro Pro Cys Ser Val Ser Pro His Thr Ser Ser Gly Ser Ser Ser Glu 165 170 175

Glu Glu Asp Ser Gly Lys Gln Ala Leu Xaa Pro Gly Leu Ser Pro Ser 180 185 190

Gln Arg Pro Gly Gly Ser Ser Ser Ala Cys Ser Arg Ser Pro Glu Glu
195 200 205

Glu Glu Glu Asp Val Leu Lys Tyr Val Arg Glu Ile Phe Phe Ser 210 215 220

<210> 205

<211> 199

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (35)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (103)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (191)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 205

Val Pro Gly Trp Pro Arg Ala Cys Ser Pro Cys Gln Ala Asp Ser Pro

1 5 10 15

Arg Ala His Pro Pro Lys Leu Arg Gly Ile Leu Arg Trp Ala Pro Val 20 25 30

Pro Leu Xaa Cys Ala Ala Leu Cys Pro Pro Leu Asp Ser Gly Met Ser 35 40 45

Met Ala Ala Cys Pro Glu Ala Pro Glu Pro Ser Phe Leu Arg Glu Val 50 55 60

Pro Ser Ser Pro Ala Ser Thr Gln Trp His Arg Pro Cys Asn Phe Arg 65 70 75 80

Gln Val Glu Ala Asn Pro Arg Lys Glu Pro Lys Asn Leu Val Trp Arg

Asp Val Ser Leu Gly Gln Xaa Ser Arg Thr Pro Arg Gly Ser Gly Leu 100 105 110

Glu Leu Val Arg Val Cys Gly Gly Gly Met Gln Arg Asp Lys Thr Val 115 120 125

Val Glu Glu Arg Val Gly Glu Glu Arg Glu Arg Glu Arg Glu Arg Glu 130 135 140

Ser Leu Gly Gly Ala Gly Lys His Gly Glu Met Arg Cys Val Tyr Val 145 150 155 160

Arg Glu Ser Val Gly Ala Pro Gly Arg Ala Gly Gly Gly Asn Gly
165 170 175

Val Asn Ser Val Gly Cys Val Arg Thr Val His Ser Gly Ser Xaa Pro 180 185 190

Pro Pro Ser Ala Gly Val Ser 195

<210> 206

<211> 174

<212> PRT

<213> Homo sapiens

142

<400> 206

Thr Arg Pro Gly Lys Glu Leu Asn Leu Val Phe Gly Leu Gln Leu Ser
1 5 10 15

Met Ala Arg Ile Gly Ser Thr Val Asn Met Asn Leu Met Gly Trp Leu 20 25 30

Tyr Ser Lys Ile Glu Ala Leu Leu Gly Ser Ala Gly His Thr Leu 35 40 45

Gly Ile Thr Leu Met Ile Gly Gly Ile Thr Cys Ile Leu Ser Leu Ile 50 55 60

Cys Ala Leu Ala Leu Ala Tyr Leu Asp Gln Arg Ala Glu Arg Ile Leu 65 70 75 80

His Lys Glu Gln Gly Lys Thr Gly Glu Val Ile Lys Leu Thr Asp Val 85 90 95

Lys Asp Phe Ser Leu Pro Leu Trp Leu Ile Phe Ile Ile Cys Val Cys
100 105 110

Tyr Tyr Val Ala Val Phe Pro Phe Ile Gly Leu Gly Lys Val Phe Phe 115 120 125

Thr Glu Lys Phe Gly Phe Ser Ser Gln Ala Ala Ser Ala Ile Asn Ser 130 135 140

Val Val Tyr Val Ile Ser Ala Pro Met Ser Pro Val Phe Gly Leu Leu 145 150 155 160

Val Asp Lys Thr Gly Lys Asn Ile Ile Trp Val Leu Cys Ala 165 170

<210> 207

<211> 31

<212> PRT

<213> Homo sapiens

<400> 207

Cys Lys Asp Leu Cys Ser Arg Val Tyr Leu Leu Thr Leu Ser Pro Leu 1 5 10 15

Leu Ser Tyr Asp Pro Ala Thr Ser His Ser Pro Arg Asn Thr Gln
20 25 30

<210> 208

<211> 369.

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (78)

<223> Xaa equals any of the naturally occurring L-amino acids

**		
l ·		

<400> 208 Ile Ile Cys Glu Cys Trp Glu Glu Glu Cys Gln Ser Cys Arg Leu Lys															
Ile 1	Ile	Суѕ	Glu	Cys 5	Trp	Glu	Glu	Glu	10	GIn	ser	СУS	AIG	15)
Ile	Thr	Gln	Pro 20	Arg	Glu	Ile	Cys	Arg 25	Met	Asp	Phe	Leu	Val 30	Leu	Phe
Leu	Phe	Туr 35	Leu	Ala	Ser	Val	Leu · 40	Met	Gly	Leu	Val	Leu 45	Ile	Суѕ	Val
Cys	Ser 50	Lys	Thr	His	Ser	Leu 55	Lys	Gly	Leu	Ala	Arg 60	Gly	Gly	Ala	Gln
Ile 65	Phe	Ser	Cys	Ile	Ile 70	Pro	Glu	Cys	Leu	Gln 75	Arg	Ala	Xaa	His	Gly 80
Leu	Leu	His	Tyr	Leu 85	Phe	His	Thr	Arg	Asn 90	His	Thr	Phe	Ile	Val 95	Leu
His	Leu	Val	Leu 100	Gln	Gly	Met	Val	Туr 105	Thr	Glu	Tyr	Thr	Trp 110	Glu	Val
		115					120					125	Leu		
	130					135					140		Leu		
145					150					155			Leu		160
His	Val	Tyr	Glu	Phe 165	Asp	Glu	Val	Met	Phe 170		Lys	Asn	Val	Arg 175	Cys
			180					185					Cys 190		
Cys	Asn	Trp 195	Cys	Val	His	Arg	Phe 200		His	His	Cys	Val 205	Trp	Val	Asn
	210					215					220	١	Tyr		
Thr 225		Thr	Ala	Ser	Ala 230		Thr	Val	Ala	11e 235	· Val	Ser	Thr	Thr	Phe 240
Leu	Val	His	Leu	Val 245		Met	. Ser	Asp	250		Gln	Glu	Thr	Tyr 255	Ile
Asp	Asp	Leu	Gly 260		Leu	His	Val	Met 265	Asp	Thr	· Val	. Ph∈	270	Ile	Glr
Туr	Leu	Phe 275		Thr	Phe	Pro	280		e Val	l Ph∈	e Met	285	ı Gly	Phe	· Vai
Val	Val 290		Ser	Phe	. Leu	Lev 295		, Gl	туз	. Lev	1 Let 300	ı Phe	e Val	. Lev	ту

			,	
		•		

144

Leu Ala Ala Thr Asn Gln Thr Thr Asn Glu Trp Tyr Arg Gly Asp Trp 305 310 315 320

Ala Trp Cys Gln Arg Cys Pro Leu Val Ala Trp Pro Pro Ser Ala Glu 325 330 335

Pro Gln Val His Arg Asn Ile His Ser His Gly Leu Arg Ser Asn Leu 340 345 350

Gln Glu Ile Phe Leu Pro Ala Phe Pro Cys His Glu Arg Lys Lys Gln 355 360 365

Glu

<210> 209

<211> 147

<212> PRT

<213> Homo sapiens

<400> 209

Leu Leu Ser Phe Lys Ile Arg Gly Leu Arg Thr Glu Asp Ala Gly Trp

1 5 10 15

Ala Gln Ser Ser Gly Gly Leu Cys Val Arg Gly Asp Ala Phe Trp
20 25 30

Met Pro Ser Ser Ser Ser Gly Leu Gly Ser Pro Ser Arg Pro Pro Ser 35 40 45

Ser Phe Leu Cys Leu Leu Leu Leu Leu Pro Pro Ala Ala Leu Ala 50 55 60

Leu Leu Phe Phe Leu Asp Phe Phe Pro Pro Arg Ala Ala Val Ser 65 70 75 80

Pro Phe Leu Pro Asp His Cys Ser Ala Arg Gln Pro Arg Val Trp Arg 85 90 95

Arg Glu Thr Leu Asn Arg Ser Ala Ser Gly Leu Gly Cys Trp Ala Arg 100 105 110

Ser Thr Glu Gln Gly Ala Val Gly Val Ala Thr Gly Thr Val Leu Asp 115 120 125

Ile Ser Leu Pro Ala Ser Cys Leu Ser Leu Trp Pro Pro Gly Pro Ser 130 135 140

Gly Gly Ile 145

<210> 210

<211> 143

<212> PRT

<213> Homo sapiens

<400> 210 Gln Leu Gly Leu Cys Leu Thr Ser Ala Ser Leu Pro Pro Ala Ser Arg 10 Cys Gly His Gln Ala Pro Leu Gly Ala Ser Asp Leu Ser Ala His His 25 Ser Ala Pro Gly Phe Ser Asp Ser Tyr Phe Thr Met Ser Cys Gln Ser 40 Ser Leu Ser Arg Ala Glu Ile Leu Gln Cys Pro Leu Val Pro Ser Val Ser Pro Pro Thr His Leu Pro Gln Gly Arg Ala Asn Lys Ser Ser Arg 70 Ala Ser Leu Pro Leu Leu Pro Gln Thr His Trp Cys Leu Phe Pro Ser Ala Arg Gly Trp Arg Arg Gly Ile Gln Ser Gly Leu Pro Pro Gly Gly 105 Ser Cys Thr Ser Pro Arg Ser Pro Pro Gln Thr Leu His Gln His Ile 120 Thr Leu Val Asn His Asn Thr Ser Tyr Trp Gln Ser Pro Ser Thr <210> 211 <211> 160 <212> PRT <213> Homo sapiens <400> 211 His Gln Pro Pro Cys Leu Leu Pro Leu Ala Val Ala Thr Arg Pro Leu 10 Trp Gly His Leu Thr Cys Leu Pro Ile Ile Leu His Leu Val Ser Val 25 Thr Leu Thr Ser Pro Cys Leu Ala Asn Gln Ala Phe Gln Gly Gln Arg 40 Ser Tyr Asn Ala Leu Trp Cys Pro Leu Phe Leu Leu Pro Thr Ser 50 Pro Lys Gly Glu Gln Thr Asn His Pro Glu Pro Ala Cys Pro Cys Phe Pro Lys Leu Thr Gly Val Phe Ser Leu Gln His Val Val Gly Ala Glu 85 Glu Phe Ser Gln Val Phe Leu Leu Val Asp Pro Val Pro Val Leu Asp 105

His Leu Leu Lys Leu Phe Thr Ser Thr Ser His Leu Leu Ile Ile Ile 115 120 125

146

Pro His Ile Gly Lys Ala Pro Ala Pro Asp Ser Leu Leu Glu Glu Leu 130 135 140

Ser Leu Ser Leu Ala Thr His Cys Lys Val Ala Val Ala Arg Phe Thr 145 150 155 160

<210> 212

<211> 157

<212> PRT

<213> Homo sapiens

<400> 212

Met Ala Ala Glu Gly Ser Arg Phe Ser Ser Gln Ser Pro Gly Leu Val 1 5 10 15

Asp Arg Gln Gly Pro Lys Cys Asp Pro Ser Arg Leu Val Ser Pro Trp
20 25 30

Gly Arg His Gly Leu Arg Ile Leu Gln Ile Gly His His Gly Arg
35 40 45

Asp Gly Gln His Glu Ala Thr His His Leu Leu Arg Val Leu Arg Ala 50 55 60

Pro Arg Val Gly Lys Ala Asp Glu Gly Ala Val Asp Ser Asp Pro Ser 65 70 75 80

Thr Pro Leu Gln Leu Lys His Glu Ala Ala His Ala Glu Asp His Ala 85 · 90 95

Gln Gln Val His Val Val Arg Arg Val Val Gln Gly Arg Val Thr
100 105 110

Phe Ala Arg Arg Gly Leu Val Pro Gln His Phe Val Arg Pro Pro Trp 115 120 125

Val Arg His Ile Val Ser Gly His Ser Glu Ser Lys Ala Arg Ser Arg 130 135 140

Leu Phe Arg Cys Arg Asn Arg Ser Phe Arg Arg Ala Ser 145 150 155

<210> 213

<211> 38

<212> PRT

<213> Homo sapiens

<400> 213

Arg Leu Val Ser Pro Trp Gly Arg His Gly Leu Arg Ile Leu Gln Ile 1 5 10 15

Gly His His Gly Arg Asp Gly Gln His Glu Ala Thr His His Leu

147

20 25 30 .

Leu Arg Val Leu Arg Ala 35

<210> 214

<211> 12

<212> PRT

<213> Homo sapiens

<400> 214

Pro Thr Asp Val Leu Lys Ile Arg Met Gln Ala Gln

<210> 215

<211> 7

<212> PRT

<213> Homo sapiens

<400> 215

Thr Tyr Glu Gln Leu Lys Arg

<210> 216

<211> 137

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (22)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (33)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (71)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 216

Arg Pro Arg Pro Ser Ala Ser Ser Leu Ala Arg Ser Ala Ser Leu Leu 1 5 10 15

Pro Ala Ala His Gly Xaa Gly Val Gly Gly Ala Gly Gly Gly Ser Ser
20 25 30

Xaa Leu Arg Ser Arg Tyr Gln Gln Leu Gln Asn Glu Glu Glu Ser Gly
35 40 45

Glu Pro Glu Gln Ala Ala Gly Asp Ala Pro Pro Pro Tyr Ser Ser Ile 50 55 60

148

Ser Ala Glu Ser Ala His Xaa Phe Asp Tyr Lys Asp Glu Ser Gly Phe 65 70 75 80

Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr Thr Leu Pro Ser Tyr Asp 85 90 95

Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr Ile Pro Leu Val Pro Gly
100 105 110

Arg Asp Glu Asp Phe Val Gly Arg Asp Asp Phe Asp Asp Ala Asp Gln
115 120 125

Leu Arg Ile Gly Asn Asp Gly Ile Phe 130 135

<210> 217

<211> 20

<212> PRT

<213> Homo sapiens

<400> 217

Arg Tyr Gln Gln Leu Gln Asn Glu Glu Glu Ser Gly Glu Pro Glu Gln 1 5 10 15

Ala Ala Gly Asp 20

<210> 218

<211> 22

<212> PRT

<213> Homo sapiens

<400> 218

Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp Phe Asp Asp Ala
1 5 10 15

Asp Gln Leu Arg Ile Gly
20

<210> 219

<211> 103

<212> PRT

<213> Homo sapiens

<400> 219

Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe 1 5 10 15

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile 20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe 35 40 45

149

Ser Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp 50 55 60 ...

Val Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn 65 70 75 80

Tyr Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg 85 90 95

Thr Arg Val Leu Phe Ile Tyr

<210> 220

<211> 198

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (29)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 220

Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu 1 5 10 15

His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln His Xaa Phe Asp Tyr

Lys Asp Glu Ser Gly Phe Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr
35 40 45

Thr Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr
50 55 60

Ile Pro Leu Val Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp 65 70 75 80

Phe Asp Asp Ala Asp Gln Leu Arg Ile Gly Asn Asp Gly Ile Phe Met
85 90 95

Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe Leu 100 105 110

Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile Ser 115 120 125

Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe Ser 130 135 140

Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val 145 150 155 160

Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
165 170 175

Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr

150

180 185 190

Arg Val Leu Phe Ile Tyr 195

<210> 221

<211> 70

<212> PRT

<213> Homo sapiens

<400> 221

Met Leu Leu His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln Phe Ser 1 5 10 15

Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val 20 25 30

Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr 35 40 45

Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr 50 55 60

Arg Val Leu Phe Ile Tyr 65 70

<210> 222

<211> 82

<212> PRT

<213> Homo sapiens

<400> 222

Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe 1 5 10 15

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile 20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe 35 40 45

Ser Thr Tyr Phe Pro Ala Phe Met Asn Ser Leu Ser Arg Ser Lys Arg 50 55 60

Thr Pro Ala Gly Ser Glu Ser Arg Cys Arg Thr Gln Arg Asn Asn His 65 70 75 80

Leu Leu

<210> 223

<211> 45

<212> PRT

<213> Homo sapiens

151

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 223

Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu 1 5 10 15

His Leu Thr Ala Ala Phe Leu Gln Arg Ala His Xaa Ile Leu Thr Thr 20 25 30

Arg Met Ser Leu Gly Phe Gln Ser Pro His Leu Thr Met
35 40 45

<210> 224

<211> 33

<212> PRT

<213> Homo sapiens

<400> 224

Met Thr Val Met Asp Pro Lys Gln Met Asn Val Ala Ala Ala Val Trp

1 10 15

Ala Val Val Ser Tyr Val Val Ala Asp Met Glu Glu Met Leu Pro Arg
20 25 30

Ser

<210> 225

<211> 189

<212> PRT

<213> Homo sapiens

<400> 225

Pro Arg Val Arg Ser Arg Glu Pro Val Ala Gly Ala Pro Gly Cys Gly
1 5 10 15

Thr Ala Gly Pro Pro Ala Met Ala Thr Leu Trp Gly Gly Leu Leu Arg
20 25 30

Leu Gly Ser Leu Leu Ser Leu Ser Cys Leu Ala Leu Ser Val Leu Leu 35 40 45

Leu Ala His Cys Gln Thr Pro Pro Ser Asp Cys Leu His Val Val Glu 50 55 60

Pro Met Pro Val Arg Gly Pro Asp Val Glu Ala Tyr Cys Leu Arg Cys 65 70 75 80

Glu Cys Lys Tyr Glu Glu Arg Ser Ser Val Thr Ile Lys Val Thr Ile 85 90 95

Ile Ile Tyr Leu Ser Ile Leu Gly Leu Leu Leu Leu Tyr Met Val Tyr
100 105 110

152

Leu Thr Leu Val Glu Pro Ile Leu Lys Arg Arg Leu Phe Gly His Ala 115 120 125

Gln Leu Ile Gln Ser Asp Asp Ile Gly Asp His Gln Pro Phe Ala 130 135 140

Asn Ala His Asp Val Leu Ala Arg Ser Arg Ser Arg Ala Asn Val Leu 145 150 155 160

Asn Lys Val Glu Tyr Ala Gln Gln Arg Trp Lys Leu Gln Val Gln Glu 165 170 175

Gln Arg Lys Ser Val Phe Asp Arg His Val Val Leu Ser 180 185

<210> 226

<211> 231

<212> PRT

<213> Homo sapiens

<400> 226

Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser

1 10 15

Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr
20 25 30

Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro 35 40 45

Ile Lys Lys Asn Ser Gly His Ile Tyr Asn Lys Asn Ile Ser Gln Lys 50 55 60

Asp Cys Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro 65 70 75 80

Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg 85 90 95

Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu 100 105 110

Gly Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro Ile 115 120 125

Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp Asp 130 135 140

Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu Ala 145 150 155 160

Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Gly Thr 165 170 175

Ala Ala Leu Glu Ala Ser Ser Pro Arg Ala Ala Lys Ser Leu Ser Leu 180 185 190

Thr Gly Met Leu Ser Ser Ala Asn Trp Gly Ile Glu Phe Lys Val Thr 195 200

Arg Lys Lys Gln Ala Asp Asn Trp Lys Gly Thr Asp Trp Val Leu Leu 220 215

Gly Phe Ile Leu Ile Pro Cys

<210> 227

<211> 456

<212> PRT

<213> Homo sapiens

<400> 227

Met Ala Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro

Leu Leu Ala Gly Leu Ala Leu Leu Gly Val Gly Pro Val Pro Ala Arg 25 20

Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly 40

Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu

Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln

Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His

Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser 105

Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser 125 115 120

Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro 135

Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile 150 155

Met Asp Phe Asn Gly Asp Leu Ile Pro Asp Ile Phe Gly Ile Thr Asn 165 170

Glu Ser Asn Gln Pro Gln Ile Leu Leu Gly Gly Asn Leu Ser Trp His 185

Pro Ala Leu Thr Thr Thr Ser Lys Met Arg Ile Pro His Ser His Ala 195 200

Phe Ile Asp Leu Thr Glu Asp Phe Thr Ala Asp Leu Phe Leu Thr Thr 220 215

154

Leu Asn Ala Thr Thr Ser Thr Phe Gln Phe Glu Ile Trp Glu Asn Leu 225 230 235 240

Asp Gly Asn Phe Ser Val Ser Thr Ile Leu Glu Lys Pro Gln Asn Met 245 250 255

Met Val Val Gly Gln Ser Ala Phe Ala Asp Phe Asp Gly Asp Gly His 260 265 270

Met Asp His Leu Leu Pro Gly Cys Glu Asp Lys Asn Cys Gln Lys Ser 275 280 285

Thr Ile Tyr Leu Val Arg Ser Gly Met Lys Gln Trp Val Pro Val Leu 290 · 295 300

Gln Asp Phe Ser Asn Lys Gly Thr Leu Trp Gly Phe Val Pro Phe Val 305 310 315 320

Asp Glu Gln Gln Pro Thr Glu Ile Pro Ile Pro Ile Thr Leu His Ile 325 330 335

Gly Asp Tyr Asn Met Asp Gly Tyr Pro Asp Ala Leu Val Ile Leu Lys 340 345 350

Asn Thr Ser Gly Ser Asn Gln Gln Ala Phe Leu Leu Glu Asn Val Pro 355 360 365

Cys Asn Asn Ala Ser Cys Glu Glu Ala Arg Arg Met Phe Lys Val Tyr 370 375 380

Trp Glu Leu Thr Asp Leu Asn Gln Ile Lys Asp Ala Met Val Ala Thr 385 390 395 400

Phe Phe Asp Ile Tyr Glu Asp Gly Ile Leu Asp Ile Val Val Leu Ser
405 410 415

Lys Gly Tyr Thr Lys Asn Asp Phe Ala Ile His Thr Leu Lys Asn Asn 420 425 430

Phe Glu Ala Asp Ala Tyr Phe Val Lys Val Ile Val Leu Ser Gly Leu 435 440 445

Cys Ser Asn Asp Cys Pro Arg Arg 450 455

<210> 228

<211> 282

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (144)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

155

<221> SITE <222> (168) <223> Xaa equals any of the naturally occurring L-amino acids <400> 228 Met Thr Lys Arg Glu Asp Gly Gly Tyr Thr Phe Thr Ala Thr Pro Glu 15 Asp Phe Pro Lys Lys His Lys Ala Pro Val Ile Asp Ile Gly Ile Ala Asn Thr Gly Lys Phe Ile Met Thr Ala Ser Ser Asp Thr Thr Val Leu Ile Trp Ser Leu Lys Gly Gln Val Leu Ser Thr Ile Asn Thr Asn Gln Met Asn Asn Thr His Ala Ala Val Ser Pro Cys Gly Arg Phe Val Ala Ser Cys Gly Phe Thr Pro Asp Val Lys Val Trp Glu Val Cys Phe Gly 90 Lys Lys Gly Glu Phe Gln Glu Val Val Arg Ala Phe Glu Leu Lys Gly 105 His Ser Ala Ala Val His Ser Phe Ala Phe Ser Asn Asp Ser Arg Arg 115 120 Met Ala Ser Val Ser Lys Asp Gly Thr Trp Lys Leu Trp Asp Thr Xaa 135 Val Glu Tyr Lys Lys Gln Asp Pro Tyr Leu Leu Lys Thr Gly Arg Phe Glu Glu Ala Ala Gly Ala Xaa Pro Cys Arg Leu Ala Leu Ser Pro 170 165 Asn Ala Gln Val Leu Ala Leu Ala Ser Gly Ser Ser Ile His Leu Tyr 180 185 Asn Thr Arg Arg Gly Glu Lys Glu Glu Cys Phe Glu Arg Val His Gly Glu Cys Ile Ala Asn Leu Ser Phe Asp Ile Thr Gly Arg Phe Leu Ala 215 Ser Cys Gly Asp Arg Ala Val Arg Leu Phe His Asn Thr Pro Gly His 230 Arg Ala Met Val Glu Glu Met Gln Gly His Leu Lys Arg Ala Ser Asn 245 250 Glu Ser Thr Arg Gln Arg Leu Gln Gln Gln Leu Thr Gln Ala Gln Glu 260 265

Thr Leu Lys Ser Leu Gly Ala Leu Lys Lys

280

275

```
<210> 229
<211> 456
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (17)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (37)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (318)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (342)
<223> Xaa equals any of the naturally occurring L-amino acids
<400> 229
Val Ile Arg His Glu Gly Ser Thr Asn Met Glu Leu Ser Gln Met Ser
Xaa Leu Met Gly Leu Ser Val Leu Leu Gly Leu Leu Ala Leu Met Ala
Thr Ala Ala Val Xaa Arg Gly Trp Leu Arg Ala Gly Glu Glu Arg Ser
         35
                                                  45
Gly Arg Pro Ala Cys Gln Lys Ala Asn Gly Phe Pro Pro Asp Lys Ser
Ser Gly Ser Lys Lys Gln Lys Gln Tyr Gln Arg Ile Arg Lys Glu Lys
Pro Gln Gln His Asn Phe Thr His Arg Leu Leu Ala Ala Leu Lys
                 85
Ser His Ser Gly Asn Ile Ser Cys Met Asp Phe Ser Ser Asn Gly Lys
                                105
Tyr Leu Ala Thr Cys Ala Asp Asp Arg Thr Ile Arg Ile Trp Ser Thr
        115
                            120
Lys Asp Phe Leu Gln Arg Glu His Arg Ser Met Arg Ala Asn Val Glu
                        135
Leu Asp His Ala Thr Leu Val Arg Phe Ser Pro Asp Cys Arg Ala Phe
145
                    150
                                        155
                                                             160
```

157

Ile Val Trp Leu Ala Asn Gly Asp Thr Leu Arg Val Phe Lys Met Thr Lys Arg Glu Asp Gly Gly Tyr Thr Phe Thr Ala Thr Pro Glu Asp Phe 185 Pro Lys Lys His Lys Ala Pro Val Ile Asp Ile Gly Ile Ala Asn Thr 200 Gly Lys Phe Ile Met Thr Ala Ser Ser Asp Thr Thr Val Leu Ile Trp Ser Leu Lys Gly Gln Val Leu Ser Thr Ile Asn Thr Asn Gln Met Asn 235 Asn Thr His Ala Ala Val Ser Pro Cys Gly Arg Phe Val Ala Ser Cys Gly Phe Thr Pro Asp Val Lys Val Trp Glu Val Cys Phe Gly Lys Lys 265 Gly Glu Phe Gln Glu Val Val Arg Ala Phe Glu Leu Lys Gly His Ser 280 Ala Ala Val His Ser Phe Ala Phe Ser Asn Asp Ser Arg Arg Met Ala 295 Ser Val Ser Lys Asp Gly Thr Trp Lys Leu Trp Asp Thr Xaa Val Glu 315 310 Tyr Lys Lys Gln Asp Pro Tyr Leu Leu Lys Thr Gly Arg Phe Glu Glu Ala Ala Gly Ala Xaa Pro Cys Arg Leu Ala Leu Ser Pro Asn Ala Gln Val Leu Ala Leu Ala Ser Gly Ser Ser Ile His Leu Tyr Asn Thr 360 355 Arg Arg Gly Glu Lys Glu Glu Cys Phe Glu Arg Val His Gly Glu Cys 375 Ile Ala Asn Leu Ser Phe Asp Ile Thr Gly Arg Phe Leu Ala Ser Cys 390 395 Gly Asp Arg Ala Val Arg Leu Phe His Asn Thr Pro Gly His Arg Ala 405 Met Val Glu Glu Met Gln Gly His Leu Lys Arg Ala Ser Asn Glu Ser Thr Arg Gln Arg Leu Gln Gln Gln Leu Thr Gln Ala Gln Glu Thr Leu 435 440 Lys Ser Leu Gly Ala Leu Lys Lys

<210> 230

<211> 363

<212> PRT

<213> Homo sapiens

<400> 230

Met Ser Val Met Val Val Arg Lys Lys Val Thr Arg Lys Trp Glu Lys

1 10 15

Leu Pro Gly Arg Asn Thr Phe Cys Cys Asp Gly Arg Val Met Met Ala 20 25 30

Arg Gln Lys Gly Ile Phe Tyr Leu Thr Leu Phe Leu Ile Leu Gly Thr 35 40 45

Cys Thr Leu Phe Phe Ala Phe Glu Cys Arg Tyr Leu Ala Val Gln Leu 50 55 60

Ser Pro Ala Ile Pro Val Phe Ala Ala Met Leu Phe Leu Phe Ser Met 65 70 75 80

Ala Thr Leu Leu Arg Thr Ser Phe Ser Asp Pro Gly Val Ile Pro Arg 85 90 95

Ala Leu Pro Asp Glu Ala Ala Phe Ile Glu Met Glu Ile Glu Ala Thr
100 105 110

Asn Gly Ala Val Pro Gln Gly Gln Arg Pro Pro Pro Arg Ile Lys Asn 115 120 125

Phe Gln Ile Asn Asn Gln Ile Val Lys Leu Lys Tyr Cys Tyr Thr Cys 130 135 140

Lys Ile Phe Arg Pro Pro Arg Ala Ser His Cys Ser Ile Cys Asp Asn 145 150 155 160

Cys Val Glu Arg Phe Asp His His Cys Pro Trp Val Gly Asn Cys Val 165 170 175

Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr Leu Phe Ile Leu Ser Leu Ser 180 185 190

Leu Leu Thr Ile Tyr Val Phe Ala Phe Asn Ile Val Tyr Val Ala Leu 195 200 205

Lys Ser Leu Lys Ile Gly Phe Leu Glu Thr Leu Lys Glu Thr Pro Gly 210 215 220

Thr Val Leu Glu Val Leu Ile Cys Phe Phe Thr Leu Trp Ser Val Val 225 230 235 240

Gly Leu Thr Gly Phe His Thr Phe Leu Val Ala Leu Asn Gln Thr Thr
245 250 255

Asn Glu Asp Ile Lys Gly Ser Trp Thr Gly Lys Asn Arg Val Gln Asn 260 265 270

Pro Tyr Ser His Gly Asn Ile Val Lys Asn Cys Cys Glu Val Leu Cys

159

275 280 285 Gly Pro Leu Pro Pro Ser Val Leu Asp Arg Arg Gly Ile Leu Pro Leu 295 Glu Glu Ser Gly Ser Arg Pro Pro Ser Thr Gln Glu Thr Ser Ser Ser 310 315 Leu Leu Pro Gln Ser Pro Ala Pro Thr Glu Leu Asn Ser Asn Glu Met 330 Pro Glu Asp Ser Ser Thr Pro Glu Glu Met Pro Pro Glu Pro Pro 345 Glu Pro Pro Gln Glu Ala Ala Glu Ala Glu Lys 360 <210> 231 <211> 184 <212> PRT <213> Homo sapiens <400> 231 Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile 25 Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser 75 70 His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys Pro Trp Val Gly Asn Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe 115 120 Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu 135 Thr Leu Lys Gly Asn Ser Trp Asn Cys Ser Arg Ser Pro His Leu Leu 145 150 Leu Tyr Thr Leu Val Arg Arg Gly Thr Asp Trp Ile Ser Tyr Phe Pro 170

Arg Gly Ser Gln Pro Asp Asn Gln

160

180

<210> 232

<211> 52

<212> PRT

<213> Homo sapiens

<400> 232

Leu His Glu Cys Leu Pro Gly Ser Ile Ser Tyr Leu His Pro Arg Thr 1 5 10 15

Pro Trp Leu Cys Leu Pro Pro Gln His Leu Ser Phe Ser Thr Phe Ser 20 25 30

Pro Pro Trp Gln Pro Ala Met Ser Pro Val Pro Gly Thr Gly Gly Pro
35 40 45

Pro Cys Gly Leu 50

<210> 233

<211> 177

<212> PRT

<213> Homo sapiens

<400> 233

Met Leu Pro Leu Leu Ile Ile Cys Leu Leu Pro Ala Ile Glu Gly Lys
1 5 10 15

Asn Cys Leu Arg Cys Trp Pro Glu Leu Ser Ala Leu Ile Asp Tyr Asp 20 25 30

Leu Gln Ile Leu Trp Val Thr Pro Gly Pro Pro Thr Glu Leu Ser Gln 35 40 45

Ser Ile His Ser Leu Phe Leu Glu Asp Asn Asn Phe Leu Lys Pro Trp 50 60

Tyr Leu Asp Arg Asp His Leu Glu Glu Glu Thr Ala Lys Phe Phe Thr 65 70 75 80

Gln Val His Gln Ala Ile Lys Thr Leu Arg Asp Asp Lys Thr Val Leu 85 90 95

Leu Glu Glu Ile Tyr Thr His Lys Asn Leu Phe Thr Glu Arg Leu Asn 100 105 110

Lys Ile Ser Asp Gly Leu Lys Glu Lys Gly Ala Pro Pro Leu His Glu 115 120 125

Cys Leu Pro Gly Ser Ile Ser Tyr Leu His Pro Arg Thr Pro Trp Leu 130 135 140

Cys Leu Pro Pro Gln His Leu Ser Phe Ser Thr Phe Ser Pro Pro Trp 145 150 155 160

161

Gln Pro Ala Met Ser Pro Val Pro Gly Thr Gly Gly Pro Pro Cys Gly 165 170 175

Leu

<210> 234

<211> 95

<212> PRT

<213> Homo sapiens

<400> 234

Pro Pro Val Pro Pro Trp Ile Ser Leu Pro Leu Thr Gly Ser Pro Pro 1 5 10 15

Arg Pro Gly Phe Val Pro Val Ser Pro Phe Cys Phe Ser Pro Met Thr
20 25 30

Asn Gly His Gln Val Leu Leu Leu Leu Leu Leu Thr Ser Ala Val Ala 35 40 45

Ala Gly Pro Trp Pro Gln Val His Ala Gly Gln Trp Gly Trp Met Cys 50 55 60

Leu Pro Pro Gly Leu Pro Ser Val Gln Ala Arg Ser Gly Leu Gly Gly 65 70 75 80

Leu Pro Gly Gly Pro Gln Trp Val Pro Gly Gly Ala Arg Gly Tyr
85 90 95

<210> 235

<211> 404

<212> PRT

<213> Homo sapiens

<400> 235

Ile Gln Gln Trp Gly Asp Ser Val Leu Gly Arg Arg Cys Arg Asp Leu

1 5 10 15

Leu Leu Gln Leu Tyr Leu Gln Arg Pro Glu Leu Arg Val Pro Val Pro 20 25 30

Glu Val Leu Leu His Ser Glu Gly Ala Ala Ser Ser Ser Val Cys Lys 35 40 45

Leu Asp Gly Leu Ile His Arg Phe Ile Thr Leu Leu Ala Asp Thr Ser 50 55 60

Asp Ser Arg Ala Leu Glu Asn Arg Gly Ala Asp Ala Ser Met Ala Cys 65 70 75 80

Arg Lys Leu Ala Val Ala His Pro Leu Leu Leu Leu Arg His Leu Pro 85 90 95

Met Ile Ala Ala Leu Leu His Gly Arg Thr His Leu Asn Phe Gln Glu 100 105 110

Phe	Arg	Gln 115	Gln	Asn	His	Leu	Ser 120	Cys	Phe	Leu	His	Val 125	Leu	Gly	Leu
Leu	Glu 130	Leu	Leu	Gln	Pro	His 135	Val	Phe	Arg	Ser	Glu 140	His	Gln	Gly	Ala
Leu 145	Trp	Asp	Cys	Leu	Leu 150	Ser	Phe	Ile	Arg	Leu 155	Leu	Leu	Asn	Tyr	Arg 160
Lys	Ser	Ser	Arg	His 165	Leu	Ala	Ala	Phe	11e 170	Asn	Lys	Phe	Val	Gln 175	Phe
Ile	His	Lys	Tyr 180	Ile	Thr	Tyr		Ala 185	Pro	Ala	Ala	Ile	Ser 190	Phe	Leu
Gln	Lys	His 195	Ala	Asp	Pro	Leu	His 200	Asp	Leu	Ser	Phe	Asp 205	Asn	Ser	Asp
Leu	Val 210	Met	Leu	Lys	Ser	Leu 215	Leu	Ala	Gly	Leu	Ser 220	Leu	Pro	Ser	Arg
Asp 225	Asp	Arg	Thr	Asp	Arg 230	Gly	Leu	Asp	Glu	G1u 235	Gly	Glu	Glu	Glu	Ser 240
Ser	Ala	Gly	Ser	Leu 245	Pro	Leu	Val	Ser	Val 250	Ser	Leu	Phe	Thr	Pro 255	Leu
Thr	Ala	Ala	Glu 260	Met	Ala	Pro	Tyr	Met 265	Lys	Arg	Leu	Ser	Arg 270	Gly	Gln
Thr	Val	Glu 275	Asp	Leu	Leu	Glu	Val 280	Leu	Ser	Asp	Ile	Asp 285	Glu	Met	Ser
Arg	Arg 290	Arg	Pro	Glu	Ile	Leu 295	Ser	Phe	Phe	Ser	Thr 300	Asn	Leu	Gln	Arg
Leu 305	Met	Ser	Ser	Ala	Glu 310	Glu	Cys	Суѕ	Arg	Asn 315	Leu	Ala	Phe	Ser	Leu 320
Ala	Leu	Arg	Ser	Met 325	Gln	Asn	Ser	Pro	Ser 330	Ile	Ala	Ala	Ala	Phe 335	Leu
Pro	Thr	Phe	Met 340	Tyr	Cys	Leu	Gly	Ser 345	Gln	Asp	Phe	Glu	Val 350	Val	Gln
Thr	Ala	Leu 355	Arg	Asn	Leu	Pro	Glu 360	Tyr	Ala	Leu	Leu	Суs 365	Gln	Glu	His
Ala	Ala 370	Val	Leu	Leu	His	Arg 375	Ala	Phe	Leu	Val	Gly 380	Met	Туг	Gly	Gln
Met 385	Asp	Pro	Ser	Ala	Gln 390	Ile	Ser	Glu	Ala	Leu 395	Arg	Ile	Leu	His	Met 400
Glu	Ala	Val	Met												

<210> 236 <211> 361

WO 00/04140

<212> PRT

<213> Homo sapiens

<400> 236

Met Leu Leu Lys His Leu Gln Arg Met Val Ser Val Pro Gln Val Lys
1 5 10 15

163

PCT/US99/15849

Ala Ser Ala Leu Lys Val Val Thr Leu Thr Ala Asn Asp Lys Thr Ser 20 25 30

Val Ser Phe Ser Ser Leu Pro Gly Gln Gly Val Ile Tyr Asn Val Ile 35 40 45

Val Trp Asp Pro Phe Leu Asn Thr Ser Ala Ala Tyr Ile Pro Ala His 50 55 60

Thr Tyr Ala Cys Ser Phe Glu Ala Gly Glu Gly Ser Cys Ala Ser Leu 65 70 75 80

Gly Arg Val Ser Ser Lys Val Phe Phe Thr Leu Phe Ala Leu Leu Gly
85 90 95

Phe Phe Ile Cys Phe Phe Gly His Arg Phe Trp Lys Thr Glu Leu Phe 100 105 110

Phe Ile Gly Phe Ile Ile Met Gly Phe Phe Phe Tyr Ile Leu Ile Thr 115 120 125

Arg Leu Thr Pro Ile Lys Tyr Asp Val Asn Leu Ile Leu Thr Ala Val 130 135 140

Thr Gly Ser Val Gly Gly Met Phe Leu Val Ala Val Trp Trp Arg Phe 145 150 155 160

Gly Ile Leu Ser Ile Cys Met Leu Cys Val Gly Leu Val Leu Gly Phe 165 170 175

Leu Ile Ser Ser Val Thr Phe Phe Thr Pro Leu Gly Asn Leu Lys Ile
. 180 185 190

Phe His Asp Asp Gly Val Phe Trp Val Thr Phe Ser Cys Ile Ala Ile 195 200 205

Leu Ile Pro Val Val Phe Met Gly Cys Leu Arg Ile Leu Asn Ile Leu 210 215 220

Thr Cys Gly Val Ile Gly Ser Tyr Ser Val Val Leu Ala Ile Asp Ser 225 230 235 240

Tyr Trp Ser Thr Ser Leu Ser Tyr Ile Thr Leu Asn Val Leu Lys Arg
245 250 255

Ala Leu Asn Lys Asp Phe His Arg Ala Phe Thr Asn Val Pro Phe Gln 260 265 270

164

Thr Asn Asp Phe Ile Ile Leu Ala Val Trp Gly Met Leu Ala Val Ser 275 280 285

Gly Ile Thr Leu Gln Ile Arg Arg Glu Arg Gly Arg Pro Phe Phe Pro 290 295 300

Pro His Pro Tyr Lys Leu Trp Lys Gln Glu Arg Glu Arg Arg Val Thr 305 310 315 320

Asn Ile Leu Asp Pro Ser Tyr His Ile Pro Pro Leu Arg Glu Arg Leu 325 330 335

Tyr Gly Arg Leu Thr Gln Ile Lys Gly Leu Phe Gln Lys Glu Gln Pro 340 345 350

Ala Gly Glu Arg Thr Pro Leu Leu Leu 355 360

<210> 237

<211> 116

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (37)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (40)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 237

Trp Ala Arg Leu Arg Gly Pro Gly Ala His Ala Arg Thr Ser Pro Gln 1 5 10 15

Pro Trp Arg Gly Pro Ser Pro Ala Gln Ala Ala Met Gly Phe Leu Gln 20 25 30

Leu Leu Val Val Xaa Val Leu Xaa Ser Glu His Arg Val Ala Gly Ala 35 40 45

Ala Glu Val Phe Gly Asn Ser Ser Glu Gly Leu Ile Glu Phe Ser Val 50 60

Gly Lys Phe Arg Tyr Phe Glu Leu Asn Arg Pro Phe Pro Glu Glu Ala 65 70 75 80

Ile Leu His Asp Ile Ser Ser Asn Val Thr Phe Leu Ile Phe Gln Ile 85 90 95

His Ser Gln Tyr Gln Asn Thr Thr Val Ser Phe Ser Pro Arg Arg Arg 100 . 105 110

Ser Pro Thr Met

165

<210> 238

<211> 166

<212> PRT

<213> Homo sapiens

<400> 238

Pro Arg Val Arg Pro Ala Ser Pro Pro Val Arg Ser Pro Ala Arg Trp

1 5 10 15 .

Gly Ser Met Ala Gly Ser Pro Leu Leu Trp Gly Pro Arg Ala Gly Gly 20 25 30

Val Gly Leu Leu Val Leu Leu Leu Gly Leu Phe Arg Pro Pro Pro 35 40 45

Ala Leu Cys Ala Arg Pro Val Lys Glu Pro Arg Gly Leu Ser Ala Ala 50 55 60

Ser Pro Pro Leu Ala Arg Leu Ala Leu Leu Ala Ala Ser Gly Gly Gln 65 70 75 80

Cys Pro Glu Val Arg Arg Gly Arg Cys Arg Pro Gly Ala Gly Ala 85 90 95

Gly Ala Ser Ala Gly Ala Glu Arg Gln Glu Arg Ala Arg Ala Glu Ala 100 105 110

Gln Arg Leu Arg Ile Ser Arg Arg Ala Ser Trp Arg Ser Cys Cys Ala 115 120 125

Ser Gly Ala Pro Pro Ala Thr Leu Ile Arg Leu Trp Ala Trp Thr Thr 130 135 140

Thr Pro Thr Arg Leu Gln Arg Ser Ser Leu Ala Leu Cys Ser Ala Pro 145 150 155 160

Ala Leu Thr Leu Pro Pro 165

<210> 239

<211> 414

<212> PRT

<213> Homo sapiens

<400> 239 ·

Pro Arg Val Arg Leu Ala Thr Pro Asn Ile Trp Asp Leu Ser Met Leu 1 5 10 15

Phe Ala Phe Ile Ser Leu Leu Val Met Leu Pro Thr Trp Trp Ile Val 20 25 30

Ser Ser Trp Leu Val Trp Gly Val Ile Leu Phe Val Tyr Leu Val Ile 35 40 45

Arg Ala Leu Arg Leu Trp Arg Thr Ala Lys Leu Gln Val Thr Leu Lys

	50					55					60				
Lys 65		Ser	Val	His	Leu 70	Glu	Asp	Met	Ala	Thr 75	Asn	Ser	Arg	Ala	Phe 80
Thr	Asn	Leu	Val	Arg 85	Lys	Ala	Leu	Arg	Leu 90	Ile	Gln	Glu	Thr	Glu 95	Val
Ile	Ser	Arg	Gly 100		Thr	Leu	Val	Ser 105	Ala	Ala	Cys	Pro	Phe 110	Asn	Lys
Ala	Gly	Gln 115	His	Pro	Ser	Gln	His 120	Leu	Ile	Gly	Leu	Arg 125	Lys	Ala	Val
Tyr	Arg 130	Thr	Leu	Arg	Ala	Asn 135	Phe	Gln	Ala	Ala	Arg 140	Leu	Ala	Thr	Leu
Туг 145	Met	Leu	Lys	Asn	туr 150	Pro	Leu	Asn	Ser	Glu 155	Ser	Asp	Asn	Val	Thr 160
Asn	Tyr	Ile	Cys	Val 165	Val	Pro	Phe	Lys	Glu 170	Leu	Gly	Leu	Gly	Leu 175	Ser
Glu	Glu	Gln	Ile 180	Ser	Glu	Glu	Glu	Ala 185	His	Asn	Phe	Thr	Asp 190	Gly	Phe
Ser	Leu	Pro 195	Ala	Leu	Lys	Val	Leu 200	Phe	Gln	Leu	Trp	Val 205	Ala	Gln	Ser
Ser	Glu 210	Phe	Phe	Arg	Arg	Leu 215	Ala	Leu	Leu	Leu	Ser 220	Thr	Ala	Asn	Ser
Pro 225	Pro	Gly	Pro	Leu	Leu 230	Thr	Pro	Ala	Leu	Leu 235	Pro	His	Arg	Ile	Leu 240
Ser	Asp	Val	Thr	Gln 245	Gly	Leu	Pro	His	Ala 250	His	Ser	Ala	Cys	Leu 255	Glu
Glu	Leu	Lys	Arg 260	Ser	Tyr	Glu	Phe	Tyr 265	Arg	Tyr	Phe	Glu	Thr 270	Gln	His
Gln	Ser	Val 275	Pro	Gln	Суѕ	Leu	Ser 280	Lys	Thr	Gln	Gln	Lys 285	Ser	Arg	Glu
Leu	Asn 290	Asn	Val	His	Thr	Ala 295	Val	Arg	Ser	Leu	Gln 300	Leu	His	Leu	Lys
Ala 305	Leu	Leu	Asn	Glu	Val 310	Ile	Ile	Leu	Glu	Asp 315	Glu	Leu	Glu	Lys	Leu 320
Val	Cys	Thr	Lys	Glu 325	Thr	Gln	Glu	Leu	Val 330	Ser	Glu	Ala	Tyr	Pro 335	Ile
Leu	Glu	Gln	Lys 340	Leu	Lys	Leu	Ile	Gln 345	Pro	His	Val	Gln	Ala 350	Ser	Asn
Asn	Cys	Trp 355	Glu	Glu	Ala	Ile	Ser 360	Gln	Val	Asp	Lys	Leu 365	Leu	Arg	Arg

167

Asn Thr Asp Lys Lys Gly Lys Pro Glu Ile Ala Cys Glu Asn Pro His 370 380

Cys Thr Val Ser Thr Phe Glu Ala Ala Tyr Ser Thr His Cys Arg Gln 385 390 395 400

Arg Ser Asn Pro Arg Gly Ala Gly Ile Arg Ser Leu Cys Arg
405
410

<210> 240

<211> 145

<212> PRT

<213> Homo sapiens

<400> 240

Ala Ala Pro His Pro Pro Leu Leu Arg Pro Leu Cys Leu Trp Cys Pro 1 5 10 15

Leu Trp Pro Ala Trp Pro Leu Arg Gly Arg Pro Arg Ser Ala Trp Lys
20 25 30

Arg Trp Pro Pro Leu Pro Val Gly Pro Ala Lys Leu Gly Cys Ser Met
35 40 45

Thr Thr Arg Gln Pro Thr Ala Val Ser Trp Pro Cys Trp Leu Met Ser 50 55 60

Ser Ser Leu Ser Thr Ala Cys Leu Ala Trp Thr Leu Thr Gly Ser Leu 65 70 75 80

Ala Arg Glu Ala Thr Arg Arg Ala Arg Ser Leu Ser Pro Thr Trp Asn 85 90 95

Cys Ser Ala Arg Gln Val Pro Pro Ser Pro Pro His Ser Gly Leu Gly
100 105 110

Arg Arg Gly Trp Ala His Cys His Leu Thr Cys Leu Leu Val Thr Gln
115 120 125

Leu Phe Arg Val Gly Arg Ile His Pro Ile Leu Ser Leu Pro Leu Val 130 135 140

Thr 145

<210> 241

<211> 72

<212> PRT

<213> Homo sapiens

<400> 241

Leu Gln Leu Ala Ser Gln Ser Ala Gly Ile Lys Gly Met Ser His Cys

1 10 15

Ala Arg Pro Thr Phe Leu Thr Leu Leu Ala Ser Cys Phe Trp Ala

168

20 25 30

Ala Ala Ile Pro Asn Arg Asn Val Ile Leu Ser Val Ser Phe Arg Pro 35 40 45

Leu His Met Gln Phe Thr Leu Ser Ile Leu Val Phe Ile Leu Arg Ile 50 55 60

Leu Ile Leu Leu Arg Ser Phe Leu 65 . 70

<210> 242

<211> 140

<212> PRT

<213> Homo sapiens

<400> 242

Met Val Leu Val Leu Arg His Pro Leu Cys Ala Arg Glu Arg Ala Phe 1 5 10 15

Arg Glu Pro Gly Arg Gly Leu Leu Thr Arg Thr Gly Gln His Asp Gly
20 . 25 30

Ala Pro Ala Val Thr Ala Val Pro Gly Pro Leu Gly Ala Val Ala Ala 35. 40 45

Ala Glu Gly Arg Arg Ser Ala Trp Gly Ala Gly Gly Ser Ser Pro Pro 50 55 60

Arg Lys Val Leu Trp Gly Asp Met Arg Gly Arg Arg Ala Gly Val Asp 65 70 75 80

Val Leu Gly Pro Ala Leu Ser Ser Glu Ala Ala Gly Ala Glu Ala Arg 85 90 95

Gly Trp Gly Met Pro Gly Met Gly Val Gly Val Gly Ala Ser Glu Thr 100 105 110

Arg Gly Ala Leu Phe Leu Gly Arg Glu Gly Val His Gly Pro Cys Pro
115 120 125

Met Asp Gly Leu Gly Pro Trp Pro Trp Gly Pro Trp 130 135 140

<210> 243

<211> 353

<212> PRT

<213> Homo sapiens

<400> 243

Met Gly Pro Ala Val Lys Met Trp Thr Asn Ala Trp Lys Gly Leu Asp 1 5 10 15

Asp Cys His Tyr Asn Gln Leu Cys Glu Asn Thr Pro Gly Gly His Arg
20 25 30

Cys	Ser	Cys 35	Pro	Arg	Gly	Tyr	Arg 40	Met	Gln	Gly	Pro	Ser 45	Leu	Pro	Cys
Leu	Asp 50	Val	Asn	Glu	Суѕ	Leu 55	Gln	Leu	Pro	Lys	Ala 60	Cys	Ala	Tyr	Glr
Cys 65	His	Asn	Leu	Gln	Gly 70	Ser	Tyr	Arg	Cys	Leu 75	Суѕ	Pro	Pro	Gly	Glr 80
Thr	Leu	Leu	Arg	Asp 85	Gly	Lys	Ala	Cys	Thr 90	Ser	Leu	Glu	Arg	Asn 95	Gly
Gln	Asn	Val	Thr 100	Thr	Val	Ser	His	Arg 105	Gly	Pro	Leu	Leu	Pro 110	Trp	Lev
Arg	Pro	Trp 115	Ala	Ser	Ile	Pro	Gly 120	Thr	Ser	Tyr	His	Ala 125	Trp	Val	Sei
Leu	Arg 130	Pro	Gly	Pro	Met	Ala 135	Leu	Ser	Ser	Val	Gly 140	Arg	Ala	Trp	Суя
Pro 145	Pro	Gly	Phe	Ile	Arg 150	Gln	Asn	Gly	Val	Cys 155	Thr	Asp	Leu	Asp	Gl:
Cys	Arg	Val	Arg	Asn 165	Leu	Суѕ	Gln	His	Ala 170	Cys	Arg	Asn	Thr	Glu 175	Gly
Ser	Tyr	Gln	Cys 180	Leu	Cys	Pro	Ala	Gly 185	Tyr	Arg	Leu	Leu	Pro 190	Ser	G13
Lys	Asn	Cys 195	Gln	Asp	Ile	Asn	Glu 200	Cys	Glu	Glu	Glu	Ser 205	Ile	Glu	Суз
Gly	Pro 210	Gly	Gln	Met	Cys	Phe 215	Asn	Thr	Arg	Gly	Ser 220	Tyr	Gln	Cys	Val
Asp 225	Thr	Pro	Cys	Pro	Ala 230	Thr	Tyr	Arg	Gln	Gly 235	Pro	Ser	Pro	Gly	Thr 240
Cys	Phe	Arg	Arg	Cys 245	Ser	Gln	Asp	Cys	Gly 250	Thr	Gly	Gly	Pro	Ser 255	Thr
Leu	Gln	Tyr	Arg 260	Leu	Leu	Pro	Leu	Pro 265	Leu	Gly	Val	Arg	Ala 270	His	His
Asp	Val	Ala 275	Arg	Leu	Thr	Ala	Phe 280	Ser	Glu	Val	Gly	Val 285	Pro	Ala	Asn
Arg	Thr 290	Glu	Leu	Ser	Met	Leu 295	Glu	Pro	Asp	Pro	Arg 300	Ser	Pro	Phe	Ala
Leu 305	Arg	Pro	Leu	Arg	Ala 310	Gly	Leu	Gly	Ala	Val 315	Tyr	Thr	Arg	Arg	Ala 320
Leu	Thr	Arg	Ala	Gly 325	Leu	Tyr	Arg	Leu	Thr 330	Val	Arg	Ala	Ala	Ala 335	Pro
Arg	His	Gln	Ser	Val	Phe	Val	Leu	Leu	Ile	Ala	Val	Ser	Pro	Tyr	Pro

WO 00/04140

170

340 345 350

Tyr

<210> 244

<211> 146

<212> PRT

<213> Homo sapiens

<400> 244

Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala 1 5 10 15

Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys
20 25 30

Phe Arg Gln Ser His Thr Trp Ser Leu Leu Gln Gly Ala Ala His Asp
35 40 45

Asp Pro Val Ala Arg Gly Leu Asp Pro Asp Gly Leu Leu Leu Asp 50 55 60

Val Val Val Asn Gly Val Val Pro Gly Arg Ala Trp Leu Thr Gln Ile 65 70 75 80

Phe Lys Cys Arg Thr Leu Lys Lys His Tyr Val Gln Thr Arg Ala Trp 85 90 95

Pro Ala Val Arg Gly Leu His Thr Ala Leu Leu Pro Gly Arg Pro Pro 100 105 110

Leu Val Pro Thr Leu Gln Pro Gln His Pro Val Gln Arg Gly Pro Gly 115 120 125

Pro Pro Ala Pro Ala Gly Ala Ala Pro Ala Gly Leu Ser Tyr Gln Leu 130 135 140

Gly Leu

145

<210> 245

<211> 638

<212> PRT

<213> Homo sapiens

<400> 245

His Ala Ser Gly Ala Phe Leu Val Val Arg Gly Glu Pro Gln Gly Ser 1 5 .10 15

Trp Gly Ser Met Thr Gly Val Ile Asn Gly Arg Lys Phe Gly Val Ala 20 25 30

Thr Leu Asn Thr Ser Val Met Gln Glu Ala His Ser Gly Val Ser Ser 35 40 45

171

Ile His Ser Ser Ile Arg His Val Pro Ala Asn Val Gly Pro Leu Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys Phe 90 Arg Gln Glu Ser His Val Glu Phe Ala Thr Gly Glu Leu Leu Thr Met Thr Gln Trp Pro Gly Val Trp Ile Pro Met Ala Ser Cys Ser Ser Thr 120 Trp Trp Ser Met Ala Leu Ser Pro Asp Ser Leu Ala Asp Ala Asp Leu 130 135 Gln Val Gln Asp Phe Glu Glu His Tyr Val Gln Thr Gly Pro Gly Gln Leu Phe Val Gly Ser Thr Gln Arg Phe Phe Gln Gly Gly Leu Pro Ser 170 Phe Leu Arg Cys Asn His Ser Ile Gln Tyr Asn Ala Ala Arg Gly Pro 185 Gln Pro Gln Leu Val Gln His Leu Arg Ala Ser Ala Ile Ser Ser Ala 200 Phe Asp Pro Glu Ala Glu Ala Leu Arg Phe Gln Leu Ala Thr Ala Leu 215 220 210 Gln Ala Glu Glu Asn Glu Val Gly Cys Pro Glu Gly Phe Glu Leu Asp . 235 Ser Gln Gly Ala Phe Cys Val Asp Val Asp Glu Cys Ala Trp Asp Ala 250 His Leu Cys Arg Glu Gly Gln Arg Cys Val Asn Leu Leu Gly Ser Tyr Arg Cys Leu Pro Asp Cys Gly Pro Gly Phe Arg Val Ala Asp Gly Ala Gly Cys Glu Asp Val Asp Glu Cys Leu Glu Gly Leu Asp Asp Cys His Tyr Asn Gln Leu Cys Glu Asn Thr Pro Gly Gly His Arg Cys Ser Cys 315 Pro Arg Gly Tyr Arg Met Gln Gly Pro Ser Leu Pro Cys Leu Asp Val Asn Glu Cys Leu Gln Leu Pro Lys Ala Cys Ala Tyr Gln Cys His Asn 345 Leu Gln Gly Ser Tyr Arg Cys Leu Cys Pro Pro Gly Gln Thr Leu Leu

355 360 365 Arg Asp Gly Lys Ala Cys Thr Ser Leu Glu Arg Asn Gly Gln Asn Val 375 Thr Thr Val Ser His Arg Gly Pro Leu Leu Pro Trp Leu Arg Pro Trp Ala Ser Ile Pro Gly Thr Ser Tyr His Ala Trp Val Ser Leu Arg Pro 410 Gly Pro Met Ala Leu Ser Ser Val Gly Arg Ala Trp Cys Pro Pro Gly 425 Phe Ile Arg Gln Asn Gly Val Cys Thr Asp Leu Asp Glu Cys Arg Val Arg Asn Leu Cys Gln His Ala Cys Arg Asn Thr Glu Gly Ser Tyr Gln 450 455 Cys Leu Cys Pro Ala Gly Tyr Arg Leu Leu Pro Ser Gly Lys Asn Cys 475 Gln Asp Ile Asn Glu Cys Glu Glu Glu Ser Ile Glu Cys Gly Pro Gly Gln Met Cys Phe Asn Thr Arg Gly Ser Tyr Gln Cys Val Asp Thr Pro 505 Cys Pro Ala Thr Tyr Arg Gln Gly Pro Ser Pro Gly Thr Cys Phe Arg 520 Arg Cys Ser Gln Asp Cys Gly Thr Gly Gly Pro Ser Thr Leu Gln Tyr 535 Arg Leu Leu Pro Leu Pro Leu Gly Val Arg Ala His His Asp Val Ala 550 555 Arg Leu Thr Ala Phe Ser Glu Val Gly Val Pro Ala Asn Arg Thr Glu 565 Leu Ser Met Leu Glu Pro Asp Pro Arg Ser Pro Phe Ala Leu Arg Pro 585 Leu Arg Ala Gly Leu Gly Ala Val Tyr Thr Arg Arg Ala Leu Thr Arg 595 Ala Gly Leu Tyr Arg Leu Thr Val Arg Ala Ala Pro Arg His Gln Ser Val Phe Val Leu Leu Ile Ala Val Ser Pro Tyr Pro Tyr 630

<210> 246

<211> 367

<212> PRT

<213> Homo sapiens

<400	0> 24	46													
Met 1	Gly	Glu	Lys	Phe 5	Leu	Leu	Leu	Ala	Met 10	Lys	Glu	Asn	His	Pro 15	Glu
Cys	Phe	Cys	Lys 20	Ile	Leu	Lys	Ile	Leu 25	His	Cys	Met	Asp	Pro 30	Gly	Glu
Trp	Leu	Pro 35	Gln	Thr	Glu	His	Суs 40	Val	His	Leu	Thr	Pro 45	Lys	Glu	Phe
Leu	Ile 50	Trp	Thr	Met	Asp	Ile 55	Ala	Ser	Asn	Glu	Arg 60	Ser	Glu	Ile	Glr
Ser 65	Val	Ala	Leu	Arg	Leu 70	Ala	Ser	Lys	Val	Ile 75	Ser	His	His	Met	Glr 80
Thr	Суѕ	Val	Glu	Asn 85	Arg	Glu	Leu	Ile	Ala 90	Ala	Glu	Leu	Lys	Gln 95	Trp
Val	Gln	Leu	Val 100	Ile	Leu	Ser	Cys	Glu 105	Asp	His	Leu	Pro	Thr 110	Glu	Ser
Arg	Leu	Ala 115	Val	Val	Glu	Val	Leu 120	Thr	Ser	Thr	Thr	Pro 125	Leu	Phe	Leu
Thr	Asn 130	Pro	His	Pro	Ile	Leu 135	Glu	Leu	Gln	Asp	Thr 140	Leu	Ala	Leu	Trp
Lys 145	Cys	Val	Leu	Thr	Leu 150	Leu	Gln	Ser	Glu	Glu 155	Gln	Ala	Val	Arg	Asp 160
Ala	Ala	Thr	Glu	Thr 165	Val	Thr	Thr	Ala	Met 170	Ser	Gln	Glu	Asn	Thr 175	Сув
Gln	Ser	Thr	Glu 180	Phe	Ala	Phe	Cys	Gln 185	Val	Asp	Ala	Ser	Ile 190	Ala	Leu
Ala	Leu	Ala 195	Leu	Ala	Val	Leu	Суs 200	Asp	Leu	Leu	Gln	Gln 205	Trp	Asp	Glr
Leu	Ala 210	Pro	Gly	Leu	Pro	11e 215	Leu	Leu	Gly	Trp	Leu 220	Leu	GJA	Glu	Ser
Asp 225	Asp	Leu	Val	Ala	Cys 230	Val	Glu	Ser	Met	His 235	Gln	Val	Glu	Glu	Asp 240
Tyr	Leu	Phe	Glu	Lys 245	Ala	Glu	Val	Asn	Phe 250	Trp	Ala	Glu	Thr	Leu 255	Ile
Phe	Val	Lys	Tyr 260	Leu	Cys	Lys	His	Leu 265	Phe	Cys	Leu	Leu	Ser 270	Lys	Ser
Gly	Trp	Arg 275	Pro	Pro	Ser	Pro	Glu 280	Met	Leu	Cys	His	Leu 285	Gln	Arg	Met
Val	Ser 290	Glu	Gln	Cys	His	Leu 295	Leu	Ser	Gln	Phe	Phe	Arg	Glu	Leu	Pro

Pro Ala Ala Glu Phe Val Lys Thr Val Glu Phe Thr Arg Leu Arg Ile 305 310 315 320

Gln Glu Glu Arg Thr Leu Ala Cys Leu Arg Leu Leu Ala Phe Leu Glu 325 330 335

Gly Lys Glu Gly Glu Asp Thr Leu Val Leu Ser Val Trp Asp Ser Tyr 340 345 350

Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu Ala Ala Cys 355 360 365

<210> 247

<211> 124

<212> PRT

<213> Homo sapiens

<400> 247

Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val 1 5 10 15

Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu 20 25 30

Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro 35 40 45

Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val 50 60

Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu 65 70 75 80

Pro Ser Cys Thr Asp Gln Cys Phe Arg Ala Lys Pro His Ser Trp Gly 85 90 95

His Phe Ser Arg Phe Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys
100 105 110

Thr Arg Asn Glu Phe Arg Leu Pro Ala Arg Ala Asp 115 120

<210> 248

<211> 674

<212> PRT

<213> Homo sapiens

<400> 248

Met Thr Gly Arg Glu Phe Phe Ser Arg Phe Pro Glu Leu Tyr Pro Phe 1 5 10 15

Leu Leu Lys Gln Leu Glu Thr Val Ala Asn Thr Val Asp Ser Asp Met 20 25 30

Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val Leu

		35					40					45			
Glu	Arg 50		Tyr	Ala	Ser	Pro 55	Met	Asp	Ğly	Thr	Ser 60		Ala	Leu	Ser
Met 65	Gly	Pro	Phe	Val	Pro 70	Phe	Ile	Met	Arg	Cys 75		His	Ser	Pro	Val 80
Tyr	His	Ser	Arg	Glu 85	Met	Ala	Ala	Arg	Ala 90	Leu	Val	Pro	Phe	Val 95	Met
Ile	Asp	His	Ile 100	Pro	Asn	Thr	Ile	Arg 105	Thr	Leu	Leu	Ser	Thr 110	Leu	Pro
Ser	Cys	Thr 115	Asp	Gln	Cys	Phe	Arg 120	Gln	Asn	His	Ile	His 125	Gly	Thr	Leu
Leu	Gln 130	Val	Phe	His	Leu	Leu 135	Gln	Ala	Tyr	Ser	Asp 140	Ser	Lys	His	Gly
Thr 145	Asn	Ser	Asp	Phe	Gln 150	His	Glu	Leu	Thr	Asp 155	Ile	Thr	Val	Cys	Thr 160
Lys	Ala	Lys	Leu	Trp 165	Leu	Ala	Lys	Arg	Gln 170	Asn	Pro	Сув	Leu	Val 175	Thr
Arg	Ala	Val	Туг 180	Ile	Asp	Ile	Leu	Phe 185	Leu	Leu	Thr	Cys	Суs 190	Leu	Asn
Arg	Ser	Ala 195	Lys	Asp	Asn	Gln	Pro 200	Val	Leu	Glu	Ser	Leu 205	Gly	Phe	Trp
Glu	Glu 210	Val	Arg	Gly	Ile	Ile 215	Ser	Gly	Ser	Glu	Leu 220	Ile	Thr	Gly	Phe
Pro 225	Trp	Ala	.Phe	Lys	Val 230	Pro	Gly	Leu	Pro	Gln 235	Tyr	Leu	Gln	Ser	Leu 240
Thr	Arg	Leu	Ala	11e 245	Ala	Ala	Val	Trp	Ala 250	Ala	Ala	Ala	Lys	Ser 255	Gly
Glu	Arg	Glu	Thr 260	Asn	Val	Pro	Ile	Ser 265	Phe	Ser	Gln	Leu	Leu 270	Glu	Ser
Ala	Phe	Pro 275	Glu	Val	Arg	Ser	Leu 280	Thr	Leu	Glu	Ala	Leu 285	Leu	Glu	Lys
Phe	Leu 290	Ala	Ala	Ala	Ser	Gly 295	Leu	Gly	Glu	Lys	Gly 300	Val	Pro	Pro	Leu
Leu 305	Cys	Asn	Met	Gly	Glu 310	Lys	Phe	Leu	Leu	Leu 315	Ala	Met	Lys	Glu	Asn 320
His	Pro	Glu	Суз	Phe 325	Cys	Lys	Ile	Leu	Lys 330	Ile	Leu	His	Cys	Met 335	Asp
Pro	Gly	Glu	Trp 340	Leu	Pro	Gln	Thr	Glu 345	His	Cys	Val	His	Leu 350	Thr	Pro

Lys	Glu	Phe 355		Ile	Trp	Thr	Met 360	Asp	Ile	Ala	Ser	Asn 365	Glu	Arg	Se
Glu	Ile 370	Gln	Ser	Val	Ala	Leu 375	Arg	Leu	Ala	Ser	Lys 380	Val	Ile	Ser	Hi
His 385	Met	Gln	Thr	Cys	Va1 390	Glu	Asn	Arg	Glu	Leu 395	Ile	Ala	Ala	Glu	Le 40
Lys	Gln	Trp	Val	Gln 405	Leu	Val	Ile	Leu	Ser 410	Cys	Glu	Asp	His	Leu 415	Pr
Thr	Glu	Ser	Arg 420	Leu	Ala	Val	Val	Glu 425	Val	Leu	Thr	Ser	Thr 430	Thr	Pr
Leu	Phe	Leu 435	Thr	Asn	Pro	His	Pro 440	Ile	Leu	Glu	Leu	Gln 445	Asp	Thr	Le
Ala	Leu 450	Trp	Lys	Cys	Val	Leu 455	Thr	Leu	Leu	Gln	Ser 460	Glu	Glu	Gln	Ala
Val 465	Arg	Asp	Ala	Ala	Thr 470	Glu	Thr	Val	Thr	Thr 475	Ala	Met	Ser	Gln	G1: 48
Asn	Thr	Суѕ	Gln	Ser 485	Thr	Glu	Phe	Ala	Phe 490	Cys	Gln	Val	Asp	Ala 495	Se
Ile	Ala	Leu	Ala 500	Leu	Ala	Leu	Ala	Va1 505	Leu	Суѕ	Asp	Leu	Leu 510	Gln	Glı
Trp	Asp	Gln 515	Leu	Ala	Pro	Gly	Leu 520	Pro	Ile	Leu	Leu	Gly 525	Trp	Leu ·	Lev
Gly	Glu 530	Ser	Asp	Asp	Leu	Val 535	Ala	Cys	Val	Glu	Ser 540	Met	His	Gln	Va:
Glu 545	Glu	Asp	Tyr	Leu	Phe 550	Glu	Lys	Ala	Glu	Va1 555	Asn	Phe	Trp	Ala	Gl: 560
Thr	Leu	Ile	Phe	Val 565	Lys	Tyr	Leu	Cys	Lys 570	His	Leu	Phe	Cys	Leu 575	Lev
Ser	Lys	Ser	Gly 580	Trp	Arg	Pro	Pro	Ser 585	Pro	Glu	Met	Leu	Суs 590	His	Leu
Gln	Arg	Met 595	Val	Ser	Glu	Gln	Cys 600	His	Leu	Leu	Ser	Gln 605	Phe	Phe	Arg
Glu	Leu 610	Pro	Pro	Ala	Ala	Glu 615	Phe	Val	Lys	Thr	Val 620	Glu	Phe	Thr	Arg
Leu 625	Arg	Ile	Gln	Glu	Glu 630	Arg	Thr	Leu	Ala	Cys 635	Leu	Arg	Leu	Leu	Ala 640
Phe	Leu	Glu	Gly	Lys 645	Glu	Gly	Glu	Asp	Thr	Leu	Val	Leu	Ser	Val 655	Trp

WO 00/04140

177

Asp Ser Tyr Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu Ala 660 665 670

Ala Cys

<210> 249

<211> 10

<212> PRT

<213> Homo sapiens

<400> 249

Ile Ile Ser Gly Ser Glu Leu Ile Thr Gly
1 5 10

<210> 250

<211> 230

<212> PRT

<213> Homo sapiens

<400> 250

Val Asp Gly Ile Asp Lys Leu Asp Ile Glu Phe Leu Gln Gln Phe Leu

1 5 10 15

Glu Thr His Ser Arg Gly Pro Arg Leu His Ser Pro Gly His Ala Ser 20 25 30

Gln Glu Ala Thr Pro Gly Ala Asn Met Ser Ser Gly Thr Glu Leu Leu 35 40 45

Trp Pro Gly Ala Ala Leu Leu Val Leu Leu Gly Val Ala Ala Ser Leu 50 55 60

Cys Val Arg Cys Ser Arg Pro Gly Ala Lys Arg Ser Glu Lys Ile Tyr 65 70 75 80

Gln Gln Arg Ser Leu Arg Glu Asp Gln Gln Ser Phe Thr Gly Ser Arg 85 90 95

Thr Tyr Ser Leu Val Gly Gln Ala Trp Pro Gly Pro Leu Ala Asp Met
100 105 110

Ala Pro Thr Arg Lys Asp Lys Leu Leu Gln Phe Tyr Pro Ser Leu Glu
115 120 125

Asp Pro Ala Ser Ser Arg Tyr Gln Asn Phe Ser Lys Gly Ser Arg His 130 135 140

Gly Ser Glu Glu Ala Tyr Ile Asp Pro Ile Ala Met Glu Tyr Tyr Asn 145 150 155 160

Trp Gly Arg Phe Ser Lys Pro Pro Glu Asp Asp Asp Ala Asn Ser Tyr

165 170 175

Glu Asn Val Leu Ile Cys Lys Gln Lys Thr Thr Glu Thr Gly Ala Gln 180 185 190 WO 00/04140

178

Gln Glu Gly Ile Gly Gly Leu Cys Arg Gly Asp Leu Ser Leu 195 200 205

Ala Leu Lys Thr Gly Pro Thr Ser Gly Leu Cys Pro Ser Ala Ser Pro 210 . 215 220

Glu Glu Asp Glu Gly Ile 225 230

<210> 251

<211> 122

<212> PRT

<213> Homo sapiens

<400> 251

Val Leu Trp Arg Glu Ala Ser Ala Leu Val Leu Ser Asn Arg Leu Ser 1 5 10 15

Ser Gly Leu Leu His Asp Leu Leu Gln Pro Ala Ile His Ser Arg
20 25 30

Leu Phe Pro Arg Arg Ser Arg Gly Leu Ser Glu Gly Glu Gly Ser Ser 35 40 45

Val Ser Leu Gln Arg Ser Arg Val Leu Ser Ala Met Lys His Val Leu 50 55 60

Asn Leu Tyr Leu Leu Gly Val Val Leu Thr Leu Leu Ser Ile Phe Val 65 70 75 80

Arg Val Met Glu Ser Leu Glu Gly Leu Leu Glu Ser Pro Ser Pro Gly 85 90 95

Thr Ser Trp Thr Thr Arg Ser Gln Leu Ala Asn Thr Glu Pro Thr Lys
100 105 110

Gly Leu Pro Asp His Pro Ser Arg Ser Met 115 120

<210> 252

<211> 129

<212> PRT

<213> Homo sapiens

' <400> 252

Tyr Thr Phe His Thr Gln Ile Phe Leu Asp Phe Pro Met Ile Phe Leu 1 5 10 15

Thr Val Leu Pro Leu Ala Phe Leu Phe Leu His Ser Gly Phe Tyr His 20 25 30

Tyr Ile Ser Phe Ser Cys Leu Phe Ser Leu Ser Leu Ala Leu Phe Phe 35 40 45

Phe Leu Asp Val Ala Thr Phe Arg Arg Pro Gly Gln Leu Phe Cys Glu

179

50 55 60

Arg Ser Val Leu Phe Asp Met Phe His Phe Gly Phe Val Ser Leu Phe 65 70 75 80

Leu His Glu Trp Ile Gln Ala Lys His Phe Trp Ala Gly Leu Phe Ile 85 90 95

Val Leu Pro Ser Asp Val Phe Phe Ser Val His His Leu Glu Ala Pro 100 105 110

Asp Gly Ser Phe Pro Asn Ile Ala Lys Leu Ser Leu Ile Ile Leu Leu 115 120 125

Arg

<210> 253

<211> 99

<212> PRT

<213> Homo sapiens

<400> 253

Gly Thr Arg Phe Pro Thr Gly Glu Thr Pro Ser Leu Gly Phe Thr Val 1 5 10 15

Thr Leu Val Leu Leu Asn Ser Leu Ala Phe Leu Leu Met Ala Val Ile 20 25 30

Tyr Thr Lys Leu Tyr Cys Asn Leu Glu Lys Glu Asp Leu Ser Glu Asn 35 40 45

Ser Gln Ser Ser Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn 50 60

Cys Ile Phe Phe Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile 65 70 75 80

Thr Ala Ile Ser Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile 85 90 95

Phe Phe Pro

<210> 254

<211> 51

<212> PRT

<213> Homo sapiens

<400> 254

Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn Cys Ile Phe Phe 1 5 10 15

Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser 20 25 30

180

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Pro Cys
35 40 45

Leu Leu Ala 50

<210> 255

<211> 259

<212> PRT

<213> Homo sapiens

<400> 255

Gly Thr Arg Phe Pro Thr Gly Glu Thr Pro Ser Leu Gly Phe Thr Val 1 5 10 15

Thr Leu Val Leu Leu Asn Ser Leu Ala Phe Leu Leu Met Ala Val Ile 20 25 30

Tyr Thr Lys Leu Tyr Cys Asn Leu Glu Lys Glu Asp Leu Ser Glu Asn 35 40 45

Ser Gln Ser Ser Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn 50 55 60

Cys Ile Phe Phe Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile 65 70 75 80

Thr Ala Ile Ser Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile 85 90 95

Phe Phe Pro Leu Pro Ala Cys Leu Asn Pro Val Leu Tyr Val Phe Phe 100 105 110

Asn Pro Lys Phe Lys Glu Asp Trp Lys Leu Leu Lys Arg Arg Val Thr 115 120 125

Lys Lys Ser Gly Ser Val Ser Val Ser Ile Ser Ser Gln Gly Gly Cys 130 135 140

Leu Glu Gln Asp Phe Tyr Tyr Asp Cys Gly Met Tyr Ser His Leu Gln 145 150 155 160

Gly Asn Leu Thr Val Cys Asp Cys Cys Glu Ser Phe Leu Leu Thr Lys 165 170 175

Pro Val Ser Cys Lys His Leu Ile Lys Ser His Ser Cys Pro Ala Leu 180 185 190

Ala Val Ala Ser Cys Gln Arg Pro Glu Gly Tyr Trp Ser Asp Cys Gly 195 200 205

Thr Gln Ser Ala His Ser Asp Tyr Ala Asp Glu Glu Asp Ser Phe Val 210 215 220

Ser Asp Ser Ser Asp Gln Val Gln Ala Cys Gly Arg Ala Cys Phe Tyr 225 230 235 240

181

Gln Ser Arg Gly Phe Pro Leu Val Arg Tyr Ala Tyr Asn Leu Pro Arg 245 250 255

Val Lys Asp

<210> 256

<211> 22

<212> PRT

<213> Homo sapiens

<400> 256

Cys Asp Cys Cys Glu Ser Phe Leu Leu Thr Lys Pro Val Ser Cys Lys
1 5 10 15

His Leu Ile Lys Ser His

<210> 257

<211> 81

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (20)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 257

Ala Leu Glu Asn Ser Gly Ser Pro Gly Leu Gln Asp Ser Ala Arg Ala 1 5 10 15

His Phe Asn Xaa Ser Leu Arg Ser Phe Ser Phe Leu Arg Asn Gln Met 20 25 30

Tyr Ile Phe Glu Leu Ser Leu Tyr Leu Glu Gly Thr Ser Phe Val Val
35 40 45

Val Leu Leu Phe Leu Leu Ile Ser Val Ser Leu Asp Ser Pro Pro Thr 50 55 60

Thr Lys Gly Trp Asp Ser Val Leu His Ile Trp Val Pro Leu Ile Val 65 70 75 80

Gln

<210> 258

<211> 77

<212> PRT

<213> Homo sapiens

<400> 258

Gly His Glu Ser Ile Cys Gly Ser Cys Arg Ser Trp Ile Tyr Phe Ser
1 10 15

182

Ile Arg Cys Arg Arg Met Arg Pro Trp Trp Ser Leu Leu Glu
20 25 30

Ala Cys Ala Thr Cys Ala Gln Thr Gly Pro Thr Arg Ser Thr Ser Cys
35 40 45

Thr Gln Glu Val Ser His Ser Ser Ser Thr Ala Tyr Pro Ala Pro Met 50 55 60

Arg Arg Cys Cys Leu Pro Ser Pro Arg Ser Cys Thr 65 70 75

<210> 259

<211> 119

<212> PRT

<213> Homo sapiens

<400> 259

Lys Arg Ala Gly Val Glu Val Gly Gly Leu Val Met Ala Leu Ala Gly 1 5 10 15

Ser Val Phe Val Leu Gly Gly Val Leu Val Leu Cys Val Glu Arg Asn 20 25 30

Gly Glu Gly Glu Met Gly Trp Pro Gln His Leu Pro Lys Ser Gln Pro 35 40 45

Leu Ser Pro Pro Val Ala Val Arg Arg Cys Ser Phe Glu Arg Ser Trp 50 55 60

Ile Asp Leu Leu Val Glu Thr Ser Ser Ser Met Val Thr Cys Arg Gln 65 70 75 80

Gln Val Gly Thr Pro Asn Gly Met Glu Gly Arg Gly Gly Pro Lys 85 90 95

Thr Thr Phe Pro Ile Arg Leu Gln Leu Ser Gly Ala Cys Ala Val Arg 100 105 110

Pro Glu Ile Gln Trp Glu Val 115

<210> 260

<211> 275

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (94)

183

<223> Xaa equals any of the naturally occurring L-amino acids

<220> ·

<221> SITE

<222> (192)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 260

Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro Phe Glu Lys Cys Met

1 5 10 . 15

Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val Thr Ile Leu Glu Ala 20 25 30

Asp Asn Arg Ile Gly Gly Arg Ile Phe Thr Tyr Arg Asp Gln Xaa Thr 35 40 45

Gly Trp Ile Gly Glu Leu Gly Ala Met Arg Met Pro Ser Ser His Arg
50 55 60

Ile Leu His Lys Leu Cys Gln Gly Leu Gly Leu Asn Leu Thr Lys Phe 65 70 75 80

Thr Gln Tyr Asp Lys Asn Thr Trp Thr Glu Val His Glu Xaa Lys Leu
85 90 95

Arg Asn Tyr Val Val Glu Lys Val Pro Glu Lys Leu Gly Tyr Ala Leu
100 . 105 110

Arg Pro Gln Glu Lys Gly His Ser Pro Glu Asp Ile Tyr Gln Met Ala 115 120 125

Leu Asn Gln Ala Leu Lys Asp Leu Lys Ala Leu Gly Cys Arg Lys Ala 130 135 140

Met Lys Lys Phe Glu Arg His Thr Leu Leu Glu Tyr Leu Leu Gly Glu 145 150 155 160

Gly Asn Leu Ser Arg Pro Ala Val Gln Leu Leu Gly Asp Val Met Ser 165 170 175

Glu Asp Gly Phe Phe Tyr Leu Ser Phe Ala Glu Ala Leu Arg Ala Xaa 180 185 190

Ser Cys Leu Ser Asp Arg Leu Gln Tyr Ser Arg Ile Val Gly Gly Trp 195 200 205

Asp Leu Leu Pro Arg Ala Leu Leu Ser Ser Leu Ser Gly Leu Val Leu 210 215 220

Leu Asn Ala Pro Val Val Ala Met Thr Gln Gly Pro His Asp Val His 225 230 235 240

Val Gln Ile Glu Thr Ser Pro Pro Ala Arg Asn Leu Lys Val Leu Lys
245 250 255

Ala Asp Val Val Leu Leu Thr Ala Ser Gly Pro Ala Val Lys Arg Ile 260 265 270 Thr Phe Ser 275

<210> 261

<211> 212

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (123)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 261

Leu Pro Arg His Met Gln Glu Ala Leu Arg Arg Leu His Tyr Val Pro
1 5 10 15

Ala Thr Lys Val Phe Leu Ser Phe Arg Arg Pro Phe Trp Arg Glu Glu 20 25 30

His Ile Glu Gly Gly His Ser Asn Thr Asp Arg Pro Ser Arg Met Ile 35 40 45

Phe Tyr Pro Pro Pro Arg Glu Gly Ala Leu Leu Leu Ala Ser Tyr Thr 50 55 60

Trp Ser Asp Ala Ala Ala Ala Phe Ala Gly Leu Ser Arg Glu Glu Ala 65 70 75 80

Leu Arg Leu Ala Leu Asp Asp Val Ala Ala Leu His Gly Pro Val Val 85 90 95

Arg Gln Leu Trp Asp Gly Thr Gly Val Val Lys Arg Trp Ala Glu Asp 100 105 110

Gln His Ser Gln Gly Gly Phe Val Val Gln Xaa Pro Ala Leu Trp Gln 115 120 125

Thr Glu Lys Asp Asp Trp Thr Val Pro Tyr Gly Arg Ile Tyr Phe Ala 130 135 140

Gly Glu His Thr Ala Tyr Pro His Gly Trp Val Glu Thr Ala Val Lys
145 150 155 160

Ser Ala Leu Arg Ala Ala Ile Lys Ile Asn Ser Arg Lys Gly Pro Ala 165 170 175

Ser Asp Thr Ala Ser Pro Glu Gly His Ala Ser Asp Met Glu Gly Gln 180 185 190

Gly His Val His Gly Val Ala Ser Ser Pro Ser His Asp Leu Ala Lys 195 200 205

Glu Glu Gly Ser 210

```
<210> 262
<211> 319
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (68)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (115)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (213)
<223> Xaa equals any of the naturally occurring L-amino acids
Met Ala Pro Leu Ala Leu His Leu Leu Val Leu Val Pro Ile Leu Leu
Ser Leu Val Ala Ser Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro
             20
                                 25
Phe Glu Lys Cys Met Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val
Thr Ile Leu Glu Ala Asp Asn Arg Ile Gly Gly Arg Ile Phe Thr Tyr
Arg Asp Gln Xaa Thr Gly Trp Ile Gly Glu Leu Gly Ala Met Arg Met
Pro Ser Ser His Arg Ile Leu His Lys Leu Cys Gln Gly Leu Gly Leu
                                     90
Asn Leu Thr Lys Phe Thr Gln Tyr Asp Lys Asn Thr Trp Thr Glu Val
            100
                                105
His Glu Xaa Lys Leu Arg Asn Tyr Val Val Glu Lys Val Pro Glu Lys
                            120
Leu Gly Tyr Ala Leu Arg Pro Gln Glu Lys Gly His Ser Pro Glu Asp
    130
                        135
Ile Tyr Gln Met Ala Leu Asn Gln Ala Leu Lys Asp Leu Lys Ala Leu
                    150
                                        155
Gly Cys Arg Lys Ala Met Lys Lys Phe Glu Arg His Thr Leu Leu Glu
                165
                                    170
Tyr Leu Leu Gly Glu Gly Asn Leu Ser Arg Pro Ala Val Gln Leu Leu
            180
                                185
                                                     190
```

186

Gly Asp Val Met Ser Glu Asp Gly Phe Phe Tyr Leu Ser Phe Ala Glu 195 Met 195 Met Ser Glu Asp 200 Per Phe Tyr Leu Ser Phe Ala Glu 205 Per Ala Ala Leu 205 Per Ala Ala Leu 210 Per Ala Ala Leu 210 Per Ala Ala Leu 210 Per Ala Ala Per Ala 210 Per Ala Ala Per Ala Per Ala Per Ala Ala Per Ala Per Ala Per Ala Ala Per Ala Per Per Per Ala Ala Ala Per Ala Ala Per Per Per Ala Ala Ala Per Ala Ala Per Per Per Ala Ala Per Ala Ala Per Per Per Ala Ala Ala Per Ala Ala Per Per Per Ala Ala Ala Per Per Per Ala Ala Ala Per Per Per Ala Ala Ala Per Ala Ala Per Per Per Ala Ala Per Ala Ala Per Per Per Ala Ala Per Ala Per Per Ala Per Al

Arg Arg Cys Gly Gly Cys Thr Thr Cys Arg Pro Pro Arg Cys Ser

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/15849

IPC(6) US CL	SSIFICATION OF SUBJECT MATTER :C12N 15/11, 15/00, 15/63; C07H 21/02, 21/04 : 536/23.1, 23.4; 435/320.1, 69.1 to International Patent Classification (IPC) or to both	national classification and IPC							
	DS SEARCHED	nandai ciassinoadon die 11 C							
	ocumentation searched (classification system followe	d by classification symbols)							
U.S. :	536/23.1, 23.4; 435/320.1, 69.1		·						
Documentat	ion searched other than minimum documentation to the	e extent that such documents are included	in the fields searched						
Sequence	Sequence searched SEQ ID NO:11, SEQ ID NO: 103, APS, STN, CAPLUS, terms nucleic acid, express, vector, pancreas islet cell tumors.								
c. Doc	UMENTS CONSIDERED TO BE RELEVANT								
Category*	Citation of document, with indication, where ap	propriate, of the relevant passages	Relevant to claim No.						
A, P	, P US 5,849,498 A (BANDMAN et al) 15 December 1998, sequence 1-10, 14-15, and listing.								
A	US 5,670,367 A (DORNER et al) 23 September 1997, especially 1-10, 14-15, and sequence listing.								
			·						
	er documents are listed in the continuation of Box C								
·	scial categories of cited documents: sument defining the general state of the art which is not considered	"T" leter document published after the interest and not in conflict with the applitude the principle or theory underlying the	lication but cited to understand						
to t	be of particular relevance	*X* document of particular relevance; th							
'L' doo	lier document published on or after the international filing date nument which may throw doubts on priority claim(s) or which is d to establish the publication date of another citation or other	considered novel or cannot be conside when the document is taken alone	ered to involve an inventive step						
spe-	cial reason (as specified)	"Y" document of particular relevance; the considered to involve an inventive combined with one or more other suc	step when the document is						
'P' doc	means being obvious to e person skilled in the art document published prior to the international filing date but later than *&* document member of the same patent family								
	priority date claimed actual completion of the international search	Date of mailing of the international sea	arch report						
30 SEPTE	MBER 1999	21 OCT 1999							
Commission Box PCT Washington	ame and mailing address of the ISA/US Commissioner of Patents and Trademarks Authorized officer								

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/15849

Box I Ob	servations where certain claims were found unsearchable (Continuation of item 1 of first sheet)
This interne	ational report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:
	Claims Nos.: Decause they relate to subject matter not required to be searched by this Authority, namely:
ا لا	Claims Nos.: Decause they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.	Claims Nos.: December they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II O	bservations where unity of invention is lacking (Continuation of item 2 of first sheet)
This Intern	ational Searching Authority found multiple inventions in this international application, as follows:
. Plca	se See Extra Sheet.
	As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3	As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
ر سا	No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Remark o	
	No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/15849

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-10, 14-15, and 21, drawn to isolated nucleic acid and expression system.

Group II, claim(s) 11-12 and 16, drawn to isolated polypeptide.

Group III, claim(s) 13, drawn to antibody.

Group IV, claim(s)17, drawn to method for preventing a medical condition.

Group V, claim(s) 18-19, drawn to method of diagnosing a disease.

Group VI, claim(s) 20, 22, drawn to method for identifying a binding partner.

Group VII, claim 23, drawn to product produced by method of claim 20.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of Group I is considered to be isolated nucleic acid and expression system.

The special technical feature of Group II is considered to be isolated polypeptide.

The special technical feature of Group III is considered to be antibody.

The special technical feature of Group IV is considered to be a method for preventing a medical condition..

The special technical feature of Group V is considered to be method of diagnosing a disease.

The special technical feature of Group VI is considered to be a method for identifying a binding partner..

The special technical feature of Group VII is considered to be product produced by method of claim 20.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as

There are 71 genes, from gene Nos 1-71.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: The 71 genes have different nucleic acid sequences and they are from different cell. Therefore they lack same special technical features.